

**Kinetic and Two-Temperature Plasma Physics of Black Hole
Accretion Disks and X-ray Coronae**

by

Amelia M. Hankla

A.B., Princeton University, 2017

M.S., University of Colorado at Boulder, 2021

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
2023

Committee Members:

Dmitri Uzdensky, Chair

Jason Dexter

Mitchell Begelman

Michael Litos

Paul Romatschke

Hankla, Amelia M. (Ph.D., Physics)

Kinetic and Two-Temperature Plasma Physics of Black Hole Accretion Disks and X-ray Coronae

Thesis directed by Prof. Dmitri Uzdensky

The accretion disks and hot X-ray coronae surrounding black holes host plasmas spanning a wide range of parameter space. The plasma can be collisional or collisionless, depending on its location relative to the black hole and properties such as the accretion rate of surrounding material onto the black hole. In these plasmas, Coulomb collisions between electrons and protons can become inefficient, resulting in a two-temperature flow. In collisionless plasmas, magnetic turbulence and reconnection can accelerate particles to Lorentz factors of 1000 or more. Modeling these processes on scales of an entire disk/corona system is difficult computationally.

In this thesis, I examine the large and small scales of black hole accretion disks and their collisionless coronae. I first study the fundamental process of how turbulence in a collisionless, magnetized coronal plasma changes in the context of an accretion disk/corona system. By driving turbulence with asymmetric energy injection, I show that the timescales for nonthermal particle acceleration depend on the injected energy's imbalance. I also propose a relativistic momentum-coupling mechanism that efficiently converts the driven electromagnetic energy into bulk kinetic energy of the plasma. Then, I demonstrate that nonthermal electrons should exist in the plunging region of a black hole. I use prescriptions from particle-in-cell simulations to build the electron distribution function within the plunging region. By ray-tracing the emission from these electrons, I show that nonthermal electrons within the plunging region create an observable power-law compatible with observations of black hole binaries in the soft spectral state. Finally, I examine two-temperature effects on the accretion disk as a whole. I probe how Coulomb collisions between protons and electrons can alter accretion disk structure, either through efficient collisions leading to disk collapse or through inefficient collisions leading to disk inflation. I contextualize these results in the framework of the disk truncation model for black hole binaries and examine the thick-to-thin disk transition as a function of accretion rate.

Dedication

To the first-generation A. Hankla, the original Dr. A. Hankla. Grandmom: you knew this moment was coming, but I still wish you could have seen it.

Acknowledgements

Graduate school is always an intense experience, made more so by the pandemic that disrupted the core middle years I spent at CU. Although I cannot possibly hope to express gratitude to everyone who deserves it, I will do my best. To those who I do not list explicitly: thank you for your advice, love, and support along the way.

I am fortunate to have worked with three amazing professors, all of whom have a unique view on science and who passed on important advice. Thank you to Mitch Begelman for fostering my love of accretion disks from the very beginning. Your vision and expertise gave me valuable perspective throughout my PhD. Thank you to Dmitri Uzdensky for your patience and thoroughness. From you, I learned the value of clear, precise scientific writing and how investigating small details can clarify or even revolutionize our understanding. Thank you to Jason Dexter for your sharp physical insight and fantastic intuition. You encouraged me to think ever more independently, regardless of what other people think.

My PhD would not have been possible without the support of both the Physics department and the Astrophysical and Planetary Sciences (APS) department. I appreciated the sense of belonging I felt in APS through colloquia, seminars, and various workshops. I also enjoyed the company of everyone in the Women and Gender Minorities in Physics (wagmip), from the workshops to casual lunches. Thank you to Jeanne Nijhowne and Gwen Dickinson, who helped me through the administrative details of getting a PhD and traveling to conferences.

My groupmates played an important role in my growth as a scientist. Many thanks to John Mehlhaff, Kai Wong, Arpi Grigorian, Michelle Athay, Coni Echibarú, Hannalore Gerling-Dunsmore, Prasun Dhang, Omar French, Greg Werner, Bhupendra Mishra, and Jake Simon for wonderful scientific discussions,

intriguing hikes, and for eating the cookies I stress-baked and the brie I stress-made. Many thanks to my other 9th-floor buddy Tyler McMaken, for everything from lunch when no one else was around, to CU Prime, to the Grand Canonical Ensemble. Thank you to Nico Scepi, who taught me how to be a rigorous scientist and an active collaborator. Thank you to Fabio Bacchini for playing such an active role as a mentor, especially encouraging me to talk to other scientists at conferences. Thank you to Alex Chen, who showed me what it was like to just drop by someone's office and discuss big ideas. Thank you to the undergraduate students who brighten my day and make me remember my passions for physics, particularly Ellie Gentry, Alex Fix, and Aaron Barrios.

Finally, my most profound thanks go to my close friends and family. To my physicist friends (Mikhail Mamaev, Diego Barbarena, Nikolay Yegotsev, Koushik Ganesan, Nanako Shitara, Thomas Bilitewski), thanks for many amazing game nights. To Bryn Grunwald, Blaise DeFranco, and Emily Kennedy — thanks for being the best friends I could ask for. Be it watching Ted Lasso, going for hikes, bonding over birds, or gaming together, you provided a much-needed respite from the sometimes all-consuming life of a graduate student. To Nermal: you have been my longest continuous friendship. Through twenty years of snuggles, you have seen me develop from an eight-year-old wanting to be a cat when I grew up to a full-fledged physicist. To Sean Muleady: thank you for being such a supportive partner as I realized that grad school is indeed difficult, for reminding me of my love of physics, for expanding my horizons, and for the small things that make us happy. I also appreciate the support of my siblings Claire and Charley and our buddy Max, our parakeets Pepper and Cat Food, as well as Sean's family: Bob, Sandy, and Daniel. Finally: Mom and Dad, without your support, I'm not sure that I would have made it through graduate school with my sanity intact. Thank you for letting us crash with you for months at a time during the pandemic, for taking care of our pets while we traveled, for spur-of-the-moment Trader Joe's treats, for delicious brunches, and for bringing us dinner during the writing of this thesis. I love you to the galactic center and back.

Contents

Chapter

1	Context and Outline	1
2	Introduction and Background	3
2.1	Plasma Regimes	3
2.1.1	Kinetic Description for Collisionless Plasmas	5
2.1.2	Fluid Description for Collisional Plasmas	6
2.1.3	Imbalanced Turbulence: from MHD to sub-Larmor Scales	6
2.2	Accretion Physics and Models	8
2.2.1	Single-temperature, Thin Accretion Disks	9
2.2.2	Two-temperature Accretion Disks	10
2.2.3	Magnetic Fields in Accretion Disks	12
2.3	Disk-Corona Models	12
2.3.1	Geometric Configurations	13
2.3.2	Plasma Physics in the Corona	14
2.4	Observational Motivation: Thermal and Nonthermal Emission	16
2.4.1	Black Hole Binary State Transitions	17
2.4.2	Spectroscopy and Timing Studies of the Disk-Corona System	19
2.5	Numerical Tools	20
2.5.1	Collisional Plasma: General Relativistic Magnetohydrodynamic Codes	22

2.5.2	Collisionless Plasmas: Particle-in-cell Codes	24
2.6	Summary of This Work	24
3	Kinetic Simulations of Imbalanced Turbulence in a Relativistic Plasma: Net Flow and Particle Acceleration	26
3.1	Introduction	27
3.2	Methods	29
3.2.1	Plasma Physical Regime	30
3.2.2	Numerical Simulations	31
3.2.3	Energy Diagnostics	35
3.3	Results	37
3.3.1	Formation of a Turbulent Cascade	37
3.3.2	Partition of the Injected Energy	40
3.3.3	Net Flow Energy and Momentum Transfer	48
3.3.4	Nonthermal Particle Acceleration	60
3.4	Conclusions	62
3.A	Dependence on Domain Size	65
4	Nonthermal Emission from the Plunging Region: a Model for the High-Energy Tail of Black Hole X-ray Binary Soft States	69
4.1	Introduction	70
4.2	Dynamical Model for the Plunging Region	72
4.2.1	Development of a Two-temperature Plasma	73
4.2.2	Nonthermal Particle Acceleration	74
4.2.3	Electron-electron Collisions	75
4.2.4	Importance of Radiative Cooling	76
4.3	Multizone Equilibrium Model	77
4.3.1	Initial Electron Distribution	77

4.3.2	Steady-state Electron Distribution	79
4.4	Results	84
4.4.1	High-spin, Highly-Magnetized Case	84
4.4.2	Parameter Space	87
4.5	Discussion	90
4.5.1	Application to Astrophysical Systems	90
4.5.2	Model Limitations	93
4.6	Conclusions	94
5	GRMHD Simulations of Accretion Disk Truncation due to Coulomb Collisions	96
5.1	Introduction	96
5.2	Methods	98
5.2.1	Electron Fluid	99
5.2.2	Electron-only Cooling Function	100
5.2.3	Injected Viscous Heating Test	101
5.3	Low Accretion Rates	101
5.3.1	Parameter Scans	102
5.3.2	Tests Injecting Viscous Heating	104
5.4	Extending to High Accretion Rates	104
5.4.1	Analytic Estimates	106
5.4.2	Numerical Attempts with Explicit Coulomb Heating	109
5.4.3	Implementing an Implicit Coulomb Solver	112
5.5	Results	119
5.5.1	Temporal Evolution	119
5.5.2	Radial Structure	122
5.5.3	Vertical Structure	125
5.5.4	Higher Target Temperature	125

5.5.5	Constant H/r	128
5.6	Discussion and Conclusions	128
5.A	Coulomb Heating Equilibrium Temperature	131
5.B	Thin Disk Scale Height	131
5.C	Implementing Isothermal Electrons	132
5.C.1	Using Fixed Electron Temperature in Coulomb Exchange	133
5.C.2	Hard-coding $T_e = 10^9$ K	135
6	Summary and Conclusions	137
	References	140
	Appendix	
A	Local Simulations of Heating Torques on a Luminous Body in an Accretion Disk	165
A.1	Introduction	166
A.2	Methods	168
A.2.1	Analytic Results	168
A.2.2	Simulations	172
A.2.3	Diagnostics	173
A.3	Results	174
A.3.1	The Linear Regime	174
A.3.2	Net Azimuthal Acceleration as a Function of Radius	176
A.3.3	Perturbation of the Surface Density	176
A.3.4	Scaling Relations	176
A.3.5	Numerical Considerations	180
A.4	Discussion	183
A.4.1	Limits of the Linear Theory	183

A.4.2	Physical Parameter Regimes	187
A.5	Conclusions	194
A.6	Acknowledgements	195
B	The Kerr Metric and Thin Accretion Disk Properties	196
B.1	Introduction	196
B.1.1	Properties of the Kerr Metric	197
B.1.2	Relativistic Thin Disk Properties	200

Tables

Table

2.1	Definitions of black hole binary spectral states. Adapted from Ref. [48]	19
3.1	Measured values of the Elsasser fields' energy ratio $r_E = \langle \overline{z_-^2} \rangle / \langle \overline{z_+^2} \rangle$ for a sampling of balance parameter ξ values. The first number gives the ratio for $N = 768$ and the second gives one standard deviation of the $N = 384$ statistical seed studies. Time averages are taken over $5.0 < tc/L < 20.0$	35
3.2	The turbulence in all simulations of balanced and imbalanced turbulence is approximately Alfvénic. The Alfvén ratio $r_A = \langle \overline{\Delta \mathcal{E}_{\text{turb}}} \rangle / \langle \overline{\Delta \mathcal{E}_{\text{mag}}} \rangle$ is approximately 1 for the largest simulations ($L/2\pi\rho_{e0} = 81.5$) for all values of the balance parameter ξ . Standard deviations are calculated from the statistical seed studies at each balance parameter.	47
3.3	Table of "equivalent" times t_{eq} in L/c where the same amount of energy density $10.3 B_0^2/8\pi$ has been injected. Values of the nonthermal particle and energy fraction at t_{eq} , shown in the second and third rows, are discussed in Section 3.3.4. Equivalent times are labelled in Fig. 3.2 as red \times 's.	61

- A.1 Summary of simulations used to calculate the scaling of net azimuthal acceleration with total injected luminosity L (Fig. A.5). All simulations have $\chi = 0.017 H^2 \Omega$, corresponding to $\lambda_c = 0.52 H$, while all other parameters (e.g. resolution, offset of the body from corotation, described in the text) are that of the fiducial simulation L1K1. This set of simulations keeps λ_c/H and x_p/λ_c constant at 0.52 and 0.187, respectively. * indicates a simulation that has density fluctuations greater than 5% of the equilibrium value and has thus entered the non-linear regime. Simulation L1K1 is often referred to as the fiducial simulation, and L100K1 as the high luminosity simulation. Measured values are presented as the average between 1 and 10 orbits plus/minus one standard deviation of the value in time. 179
- A.2 Summary of simulations used to calculate the scaling of net azimuthal acceleration with conductivity χ (Fig. A.6). All simulations have $L = 1.9 \times 10^{-4} P_0 \Omega H^3$ ($\ell_v = 1.42 P_0 \Omega$), while all other parameters (e.g. resolution, offset of the body from corotation) are that of the fiducial simulation L1K1. The value of λ/H is easily read off; the value of x_p/λ_c is obtained by noting that $x_p = 0.097 H$. * indicates a simulation that has density fluctuations greater than 5% of the equilibrium value and has thus entered the non-linear regime. Simulation L1K1 is often referred to as the fiducial simulation, and L1K10 as the high conductivity simulation. Measured values are presented as the average between 1 and 10 orbits plus/minus one standard deviation of the value in time. 181
- A.3 Summary of important values for different simulations. Ref. [185] studies the cold thermal torque (no luminosity), while Ref. [186] studies the cold and heating torques in a semi-global simulation. The values in the table were obtained either from Sec. 5.3.2 in [187] (†), by private communication (*; in particular $H/R = 0.036$), or by calculation from parameters found in the respective work (§). Unmarked values are straightforward calculations from other values. Physical values for the fiducial simulation L1K1 were calculated assuming a body orbiting at 5.2 AU (the same as [186]) around a solar-mass central body. Here $\lambda_c \equiv k_c^{-1}$ rather than as in Eq. A.4 in accordance with Ref. [187]’s definition. 192

Figures

Figure

- 2.1 Sampling of plasma systems in number density and temperature parameter space. Astrophysical plasmas (stars excepted) generally have much lower densities than naturally-occurring plasmas on Earth, which in turn are generally sparser than laboratory-created plasmas. Text color indicates the collisionality of the system, which depends on system size in addition to density and temperature: red for collisionless and black for collisional. System parameters were estimated from Refs. [8–11]. 4
- 2.2 Possible configurations of the disk/corona system. The thermal disk emission (red) is upscattered by hot electrons in the corona. The resulting high-energy photons (blue) reflect off the disk, which then emits “reprocessed” light (green). Adapted from Ref. [47]. 13
- 2.3 a) An illustration of the disk truncation model for BHB state transitions. At high luminosities and accretion rates, a thin accretion disk reaches all the way to the ISCO. At smaller luminosities, the thin disk truncates and a thick disk populates the inner region. Ch. 5 investigates the intermediate regime. Figure adapted from Ref. [76]. b) Schematic of a possible mechanism for particle acceleration in a sandwich-type corona (blue) sitting above a thin disk (green) threaded by large-scale magnetic fields (gray). Alfvén waves (red) launch predominantly upward into the corona and then reflect. Green arrows show the dominant direction of energy flux that starts the turbulent cascade. Ch. 3 investigates heating and particle acceleration in this scenario. 15

- 2.4 Illustration of a) plasma heating/energization and b) nonthermal particle acceleration. Energy injected into a system with an initial distribution function in blue can end with a final distribution function shown in orange. Electrons and protons do not need to experience the same effect from each process. Here I assume an isotropic particle distribution function and have normalized to the particle number density. 15
- 2.5 Black hole binaries undergo outbursts that significantly change their X-ray luminosity and spectral shape. Panel a shows the counts per second over time (Modified Julian Date; MJD) detected from the binary XTE J1859+226 by the All-sky Monitor (ASM) aboard the Rossi X-ray Timing Explorer (RXTE), which dramatically increase over the course of an outburst. Panel b shows spectra recorded from observations of the binary Cyg X-1 during two distinct X-ray states: the high/soft state in red and the low/hard state in blue. Panel c demonstrates how model fits to spectral data from panel a characterize the time evolution of the binary's spectral states. Adapted from Ref. [48] and Ref. [78]. 18

2.6 Interpreting observational data to extract black hole spin relies on fitting several components of the reflection spectrum, which is assumed to come from hot, presumably thermal electrons located somewhere around the black hole. Relativistic effects, which depend sensitively on how close to the black hole the radiation travels, blur the spectrum further and form the basis for iron-line-based spin measurements. a) Actual observations from the Suzaku, with different detectors shown in orange, red, blue, and pink. Model component fits are labeled and shown in black. The bottom panel shows the ratio of the data to the model. Adapted from Ref. [93]. b) Example reflection spectrum before (red) and after (blue/green) relativistic blurring. The green line shows the relativistically blurred spectrum for a non-spinning black hole, while the blue line shows the spectrum for a black hole spin of $a = 0.998$. From Ref. [47]. c) Current models can drastically overestimate black hole spin, motivating the need for better understanding of the plasma physics in the corona. Data from a 3D MHD simulation, scaled to have an input spin a_{real} , is fed through a model that measures spin through the fluorescent iron broadening. The model's extracted spin a_{fit} is shown on the y-axis. The dotted black line shows the perfect recovery case, while colored lines show the dependence on disk scale height at the ISCO. From bottom to top: $h/r_g \in 0.01, 0.25, 0.5, 1.0$. From Ref. [94]. 21

- 3.1 A turbulent cascade forms for all balance parameters. a) The magnetic energy spectra $P_{\text{mag}}(k_{\perp})$ for $L/2\pi\rho_{e0} = 81.5$ simulations of varying balance parameter averaged between times $8.8 < tc/L < 9.9$ (comprising five outputs) show an inertial range between $k_{\perp}\rho_e(t) \sim 0.1$ and 1.0 . A break in the spectrum at $k_{\perp}\rho_e(t) \sim 1.0$ indicates the onset of kinetic effects. b) When compensated by k_{\perp}^2 , the spectra for the balanced $\xi = 0.75$ and 1.0 cases are slightly steeper than $\propto k_{\perp}^{-2}$, whereas the imbalanced case $\xi = 0.0$ is slightly flatter. The Elsasser fields' spectra, shown in dash-dot red lines for $\xi = 0$, exhibit slightly different slopes, with the stronger field (z_+ , top line) being slightly steeper than the weaker field (z_- , bottom line). In both panels, shaded lines show one temporal standard deviation about the mean. Black dashed lines show the scaling $k_{\perp}^{-5/3}$; black dotted lines show k_{\perp}^{-2} . Gray lines show the $L/2\pi\rho_{e0} = 164$ balanced simulation's magnetic energy spectrum, taken at $t = 8.9 L/c$ 38
- 3.2 The amount of energy injected into a simulation depends on its balance parameter. The simulations of more balanced turbulence (purple and blue) have more injected energy than the simulations of less balanced turbulence (yellow and green). Red \times 's indicate the "equivalent" times where the same amount of energy has been injected for each simulation (see Table 3.3), which all have $L/2\pi\rho_{e0} = 81.5$ 41

- 3.3 Energy partition into electromagnetic, turbulent kinetic, internal, and net flow energy depends on balance parameter. Left column: each type of energy density evolved over time, normalized to the constant value of the initial magnetic energy density $B_0^2/8\pi$. The turbulent electromagnetic (a) and kinetic (b) energy densities reach a constant value whereas the internal (c) and net flow (d) energy densities increase over time. Right column: the change in each type of energy density evolved over time, normalized to the total amount of injected energy density $\mathcal{E}_{\text{inj}}(t)$. Summing over the four panels on the right for each simulation adds to 1. Turbulent electromagnetic (e) and kinetic (f) energy efficiencies decay as $\propto t^{-1}$, whereas internal (g) and net flow (h) energy efficiencies saturate at a constant fraction of the injected energy. Note that the net flow energy (h) has a different vertical axis. Colors and markers indicate balance parameter. These simulations all have $L/2\pi\rho_{e0} = 81.5$ 42
- 3.4 Time evolution of the energy partition for the balanced case ($\xi = 1.0$; left) and most imbalanced case ($\xi = 0.0$; right). Both show a decay in turbulent electromagnetic and kinetic energies and a saturation of internal and net flow energy densities; black dashed lines shows fits to $A + B/t$, with A and B constants. The imbalanced case has net flow energy density about an order of magnitude higher than the balanced case and correspondingly lower internal energy density, whereas the turbulent electromagnetic and kinetic energy densities are comparable for both balance parameters. 45
- 3.5 Trends of the turbulent magnetic and kinetic energy densities with balance parameter. Quantities are time-averaged from $10 < tc/L < 20$. The largest domain size $L/2\pi\rho_{e0} = 81.5$ (filled markers) shows a linear trend with balance parameter for the turbulent magnetic (a) and kinetic (b) energy densities, respectively. The statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). The dashed lines show linear fits. Colors and markers are the same as in Fig. 3.3. 45

- 3.6 The injection efficiency η_{inj} (a; Equation 3.18) and cascade time τ_{casc} (b; Equation 3.19) depend on balance parameter. The largest domain size $L/2\pi\rho_{e0} = 81.5$ is shown with filled markers and the statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). Colors and markers are the same as in Fig. 3.3. 49
- 3.7 The net flow energy efficiency decreases with increasing balance parameter. The plotted values are volume-averaged and time-averaged from $10 < tc/L < 20$. The largest domain size $L/2\pi\rho_{e0} = 81.5$ is shown with filled markers and the statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). Note that the outliers for $\xi = 0.5, 0.75$, and 1.0 with net flow energy efficiencies a factor of 2 higher than the rest of the seed study were run with the same random seed. Colors and markers are the same as in Fig. 3.3. 52
- 3.8 Even turbulence that is balanced as a whole has spatial and temporal pockets of locally imbalanced turbulence. Slices of the Poynting flux $(c/4\pi)\mathbf{E} \times \mathbf{B}$ in the z -direction, taken at the plane $z = 0$ at time $t = 16.1 L/c$ and normalized to $(c/4\pi)B_0^2$ for balanced turbulence ($\xi = 1.0$; left) and most imbalanced turbulence ($\xi = 0.0$; right) show variation in the sign of Poynting flux throughout the domain. 53
- 3.9 The average parallel Poynting flux is approximately constant in time, whereas the z -momentum of the plasma increases linearly in time. a) The time evolution of the volume-averaged Poynting flux $(c/4\pi)\langle \mathbf{E} \times \mathbf{B} \rangle$ in the z -direction, normalized to $(c/4\pi)B_0^2$, shows fluctuations around some mean value; time-averaging the curves over $10 < tc/L < 20$ shows a quadratic dependence on balance parameter (b). The time evolution (c) of the parallel plasma momentum shows an increase in time. The ratio $\langle \mathcal{P}_{z,\text{tot}} c^2 \rangle / (\langle S_z \rangle tv_{A0}/L)$, shown in (d), is of order unity for all values of balance parameter. Black dash-dot lines show a quadratic fit; the dotted line is a quadratic fit without the outlier seed. Dashed black lines indicate zero. Colors and markers are the same as in Fig. 3.3. 54

- 3.10 The parallel electromagnetic momentum generates and maintains the parallel momentum of the plasma. In these plots, each point corresponds to the instantaneous values of the volume-averaged plasma momentum along the background magnetic field and the volume-averaged parallel electromagnetic momentum at a given simulation time. Three periods of time are shown with linear fits of slopes 2.2, 3.9, and 7.4 for (a) $5 < tc/L < 10$, (b) $10 < tc/L < 15$, and (c) $15 < tc/L < 20$, respectively. Here, $g_0 = B_0^2/(4\pi c)$ is a typical value of momentum density. Colors and markers are the same as in Fig. 3.3. 55
- 3.11 The net plasma velocity along the background magnetic field depends on balance parameter.
- a) The time evolution of the volume-averaged velocity v_{net} (Equation 3.21) shows values that fluctuate in time around some mean value that depends on ξ ; time-averaging the curves over $10 < tc/L < 20$ shows a dependence on balance parameter (b). The largest domain size $L/2\pi\rho_{e0} = 81.5$ is shown with filled markers and the statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). Colors and markers are the same as in Fig. 3.3. Black dashed lines show zero. 55
- 3.12 Particle acceleration occurs for all values of balance parameter. a) The distribution function of the most imbalanced case ($\xi = 0.0$) becomes shallower in time from an initial Maxwell-Jüttner distribution (purple) to a Maxwell-Jüttner distribution plus a hard power-law component at later times (yellow). b) Spectra taken at the same time $t = 8.0 L/c$ for different balance parameters show different peak energies but similar power-law components. The dashed black line shows a Maxwell-Jüttner fit to the $\xi = 0.0$ case. The vertical green dash-dot line shows the mean Lorentz factor $\langle\gamma\rangle$ extracted from this fit. The vertical green dotted line shows the maximum energy γ_{max} . The dash-dot black line in panel (a) shows the power law $\gamma^{-2.7}$, while the dotted black line in panel (b) shows γ^{-3} . Colors are the same as in Fig. 3.3. 58

- 3.13 At equivalent times, the power laws of imbalanced turbulence are slightly flatter/harder than balanced turbulence. a) The particle energy spectra at equivalent times (see Table 3.3) show a similar mean energy and similar power laws until $\gamma \sim 10^4$. The dashed black line shows a Maxwell-Jüttner fit to the $\xi = 0.0$ case. b) Compensating by γ^3 reveals that more imbalanced turbulence ($\xi = 0.0, 0.25,$ and 0.5) has flatter power laws than the more balanced turbulence with $\xi = 0.75$ or 1.0 . The vertical green dash-dot line shows the mean Lorentz factor $\langle \gamma \rangle$ extracted from the Maxwell-Jüttner fit. The vertical green dotted line shows the maximum energy γ_{\max} . The black dash-dot line in panel (b) shows the spectrum compensated to γ^3 , while the dotted black line shows the power law γ^{-3+3} (a constant). Colors are the same as in Fig. 3.3. 59
- 3.14 The partition of plasma energy $\langle \mathcal{E}_{\text{pl}} \rangle$ into thermal and nonthermal components shows a moderate increase with the balance parameter at any given time. Both the fraction of particles with nonthermal energies (a) and the fraction of total plasma energy density $\langle \mathcal{E}_{\text{pl}} \rangle$ contained in such particles (b) are calculated by fitting a thermal Maxwell-Jüttner function to the low- and medium-energy particle distribution at each time and subtracting the fit from the total particle distribution. Colors are the same as in Fig. 3.3. 63
- 3.15 The injection efficiency and Poynting flux along the background magnetic field depend weakly on simulation domain size within statistical variation. When averaged from $t = 5 - 14 L/c$, the injection efficiency (a) and Poynting flux along the background magnetic field (b) are shown as a function of $L/2\pi\rho_{e0}$. Both the balanced ($\xi = 1.0$; purple circles) and most imbalanced values ($\xi = 0$; yellow triangles) are mostly within statistical variation of the $L/2\pi\rho_{e0} = 40.7$ sample of 8 random seeds. The black dashed line indicates zero. 66
- 3.16 The magnetic and turbulent kinetic densities are weakly dependent on simulation domain size. When averaged from $t = 5 - 14 L/c$, the magnetic energy density (a) and turbulent kinetic energy density (b) are shown as a function of $L/2\pi\rho_{e0}$. Both the balanced ($\xi = 1.0$; purple circles) and most imbalanced ($\xi = 0$; yellow triangles) values are mostly within statistical variation of the $L/2\pi\rho_{e0} = 40.7$ sample of 8 random seeds. 67

- 4.1 Properties of the plunging region calculated from the Ref. [51] background for fiducial parameters $a = 0.95$ and $\Delta\epsilon = 1.04$. From left to right: radial magnetic field strength B_r (left axis) and the fluid's Lorentz factor (right axis), dimensionless ion temperature $\theta_i = k_B T_i / m_p c^2$ (left axis) and ratio of gas-to-magnetic pressure $\beta_i = 8\pi n_0 k_B T_i / B_r^2$ (right axis), electron scattering optical depth τ_{es} , and cold ion magnetization σ_i . Here, $B_0 = 10^8$ G. The plunging region extends from the ISCO at $1.94r_g$ to the event horizon at $1.31r_g$ in the midplane. 73
- 4.2 Timescales in the plunging region show that all physical processes except for electron-ion collisions are fast compared to the infall time for a wide range of accretion efficiencies $\Delta\epsilon$, consistent with the hierarchy $t_{\text{accel}} \ll t_{\text{cool}}(\gamma) \ll t_{\text{coll}}^{ee}(\gamma) \ll t_{\text{infall}} \lesssim t_{\text{therm}}^{ei}$. Timescales are shown at a radius $r = 1.58r_g$ for black hole spin $a = 0.95$. These timescales are calculated for thermal ion and hybrid electron distributions determined by the steady-state model described in Section 4.3. 78
- 4.3 Radial profiles of the electron distribution function model properties for fiducial parameters $a = 0.95$ and $\Delta\epsilon = 1.04$. From left to right: the high-energy Lorentz factor cut-off of the power law γ_2 , the electron power-law index, the thermal electron temperature $\theta_e = k_B T_e / m_e c^2$, and the fraction of the total cooling rate Q_- from the power-law and thermal component. Vertical dashed line shows the location of the half-light radius. 79
- 4.4 The effective power-law index p_{eff} is calculated from the initial power-law distribution $f_{\text{PL},0}$ by assuming that both power laws have the same value at the high-energy cut-off γ_2 (Equation 4.17). This sample electron distribution was calculated for fiducial parameters close to the event horizon at a radius $r = 1.31 r_g$, where $p = 3.3$ (purple dashed line) and $p_{\text{eff}} = 0.6$ (blue solid line). For radii closer to the ISCO, the difference between $f_{\text{PL},0}$ and f_{PL} is not so pronounced. The thermal distribution is also shown for reference (green dotted line). 82

- 4.5 Spectrum for the fiducial model parameters $a = 0.95$ and $F_{\theta\phi} = 6.0$ at an inclination of 60° . The thin disk emission (dashed orange line) peaks around a few keV. The nonthermal electrons' radiation (solid blue line) extends from 10 eV to 0.1 GeV, although emission beyond 1 MeV (gray region) may be impacted by pair creation. The thermal electrons in the plunging region create a small excess at around 10 keV (dotted green line). 88
- 4.6 Portion of the fiducial model's observed frequency-integrated luminosity originating between the event horizon and a radius r for an inclination angle of 60° . Emission from the plunging region's nonthermal electrons (blue) dominates over the plunging region's thermal electrons (green). 10% of the luminosity comes from $r < r_{1/10} = 1.66r_g$ (dotted line), while 50% comes from $r < r_{1/2} = 1.77r_g$ (dashed line). Vertical dash-dot line shows the decoupling radius. 89
- 4.7 Frequency-integrated luminosity in X-rays ($1 \text{ keV} < \nu < 1 \text{ MeV}$) in predicted power law compared to the thermal blackbody disk for different model parameters, assuming an inclination angle $i = 60^\circ$. Models where nonthermal electron emission dominates over thermal electron emission, satisfying $L_{\text{PL}} > L_{\text{MJ}}$, are marked with large black circles. 91
- 4.8 Power-law fraction (Equation 4.25) as a function of inclination angle. Calculated from the most strongly magnetized plunging region parameters such that nonthermal plunging region emission dominates thermal plunging region emission. 92
- 5.1 Changing a variety of parameters did not eliminate the problem of the electron temperature being much higher than the target at inner radii. The panels show the dependence of density-weighted shell-averages time-averaged from 1500 to 2500 r_g/c on a) electron cooling time, b) electron adiabatic index, c) σ_{cut} , and d) s . These parameters are defined in Eq. 5.9 or Sec. 5.2.2. Dotted vertical lines show the location of the ISCO for this black hole spin of $a = 0.9375$ 103

5.2	Test 2D GRMHD simulations at low accretion rate ($\dot{m} \sim 10^{-7}$ showing the impact of prescribing viscous heating as opposed to the MHD heating method. Data are density-weighted shell averages, averaged over times in the interval $(3500, 4500)r_g/c$. Dotted vertical line shows the ISCO; dashed horizontal line shows the target temperature for the electron cooling function.	105
5.3	Coulomb heating rate as a function of a) electron temperature and b) proton temperature, calculated using Eq. 5.13.	107
5.4	Plots of the Coulomb quality factor (Eq. 5.14) at low accretion rates (panel a) and high accretion rates (panel b) over proton temperature. Panel c fixes the electron temperature at a target value of 10^9 K and shows how the Coulomb quality factor depends on mass density.	108
5.5	Illustration of unphysical behavior due to explicit evolution of Coulomb heating in the high-density regime. Plots show density-weighted shell averages taken from high-resolution GRMHD simulations, averaged over $13,000$ to $15,000r_g/c$	110
5.6	The unphysical Coulomb exchange rate that breaks the explicit evolution happens predominantly in the midplane of the accretion disk, where densities are highest. These slices are taken at $\phi = 0$ at a snapshot in time from high-resolution 3D GRMHD simulations.	111
5.7	Gas box test for Coulomb thermalization in the high-density, single-temperature regime where the explicit solver (left) breaks and sets T_p/T_e to its ceiling value, compared to the implicit solver (right) that is meant for high densities.	114
5.8	Gas box test for Coulomb heating in the single-temperature regime. The electrons and protons should cool down to the same temperature to the target, but the operator-splitting implementation breaks down (left) whereas the non-operator-splitting implementation behaves as expected (right).	116
5.9	Two-species thermalization tests for gas box in flat space-time.	117

- 5.10 2D GRMHD tests of the Coulomb heating implementation at low accretion rate M7, where it should not make a difference to use an implicit or explicit evolution. Plots show density-weighted shell averages time-averaged from $4500 < tc/r_g < 5000$. Blue lines show the explicit evolution, which holds for this low density limit, while orange and green show the two implicit methods, one that uses operator-splitting (green) and one that does not (orange). Compare to the high-density regime in Fig. 5.11. 118
- 5.11 2D GRMHD tests of the Coulomb heating implementation at high accretion rate M3; density-weighted shell averages time-averaged from $4500 < tc/r_g < 5000$. Blue lines show the explicit evolution, which breaks down for this high density limit, as seen in particularly in the ratio T_p/T_e . Orange and green lines show the two different implicit method implementations. Because the operator split method (green) shows signs of breaking down in the single-temperature limit, we prefer the non-operator split method (orange). Compare to the low-density regime in Fig. 5.10. 120
- 5.12 Density-weighted shell averages at $r = 2.04r_g$ (left) and $r = 10.0r_g$ (right) plotted over time to demonstrate the lack of steady-state. The speed of equilibration depends both on the radial location and the accretion rate. The target temperature 10^9 K is shown in the top row as a dotted horizontal line. Dotted lines in the bottom row show the minimum and maximum ratios for T_p/T_e and the single-temperature case of $T_p/T_e = 1$ 121
- 5.13 Density-weighted shell averages of relevant proton and electron timescales for three different accretion rates. Timescales are normalized to the dynamical time $t_{\text{dyn}} = 1/\Omega$. Here, the black vertical dotted line shows the location of the ISCO. The black horizontal dotted lines show where the timescales equal the dynamical time. 123
- 5.14 Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. Dotted vertical black line shows the location of the ISCO. Dotted vertical colored lines show the location of the infall equilibrium radius r_{eq} where $t = r_{eq}/|v^r|$. Dotted black lines in the T_p/T_e and $Q_{\text{coul}}/Q_{\text{visc}}^e$ plots shows where the ratio equals 1. Dotted black line in the H/r plot shows the theoretical thin disk scaling for $T_p = T_e = T_{\text{target}}$ 124

- 5.15 Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. Dotted vertical black line shows the location of the ISCO. Dotted vertical colored lines show the location of the infall equilibrium radius r_{eq} where $t = r_{eq}/|v^r|$ 126
- 5.16 Azimuthally-averaged vertical slices for three accretion rates, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. The solid contour shows $\sigma = 1$. The dashed contours show $\rho = 1$ and 10. The dotted contours show $\beta = 1$ 127
- 5.17 Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. Simulations have an accretion rate that is about the same, $\dot{m} \sim 10^{-3} \dot{m}_{\text{Edd}}$ 129
- 5.18 Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. 130
- 5.19 Density-weighted shell averages from 2D GRMHD torus tests with isothermal electrons, time-averaged over $4500 < tc/r_g < 5000$. This implementation sets $T_e = 10^9$ K only while calculating the Coulomb energy exchange rate. It does not fix the problems at high accretion rate, which are likely due to using an explicit evolution for the Coulomb heating (Ch. 5.4.3). 134
- 5.20 Density-weighted shell averages from 2D GRMHD torus tests, time-averaged over $4500 < tc/r_g < 5000$. This implementation sets $T_e = 10^9$ K only while calculating the Coulomb energy exchange rate. Because Q_{coul} is extremely large with seemingly no impact on the disk structure and H/r , something must be wrong with this implementation. 136

- A.1 Illustration of the physics leading to a heating torque (the gravitational potential of the body is neglected). Heat diffusing away from a luminous, disk-embedded body, is sheared by the sub-Keplerian disk flow, forming hot low density lobes. These lobes are asymmetric interior to / exterior to the body, because the body is displaced from exact co-rotation due to the presence of a pressure gradient in the disk gas. The gravitational back-reaction from the heated lobes exerts a positive torque on the orbiting body, as illustrated by the vectors showing the *perturbed* force in the azimuthal direction, i.e. the difference between the force on a luminous massless body and the force on a non-luminous massless body. Other components of the force (e.g. in the radial direction) are not drawn because they are opposite and equal and thus do not contribute to the total torque. 169
- A.2 Slices in the $z = 0$ plane of simulations with $\chi = 0.017 H^2\Omega$ at $t = 5.0$ orbits. Top row: density perturbation as a percentage of initial (equilibrium) condition. Bottom row: perturbation in pressure as a percentage of initial (equilibrium) condition. Left column: $\ell_v = 1.42 P_0\Omega$ (fiducial simulation L1K1, linear regime). Right column: $\ell_v = 142 P_0\Omega$ (high luminosity simulation L100K1, non-linear regime). Note that (though similar in shape) the perturbations in L100K1 are larger in magnitude than the perturbations in L1K1. Velocity streamlines with the background shear profile subtracted are plotted in the lower panels, normalized to their maximum values for ease of comparison between L1K1 and L100K1. The distance from co-rotation x_p , the characteristic wavelength λ_c , and the direction of shear are marked with black arrows. 175

- A.3 Snapshot of gravitational acceleration on the body in the fiducial simulation L1K1 as a function of distance away from the body at $t = 5.0$ orbits. The top panel plots both the one-sided forces due to gas in front of ($y > 0$) and behind ($y < 0$) the body, as well as the sum of the forces (dashed green line; right scale). Vertical dashed lines show the size of the luminosity injection radius $r_{\text{rad}} = 0.04 H$ and half the characteristic wavelength $\lambda_c = 0.52 H$. The bottom panel shows the difference between the initial condition (which has net force equal to zero but one-sided forces on the order of the top panel's vertical values) and their values at five orbits. For reference the sum over all radii (with value $3.03 \times 10^{-3} H\Omega^2$) is plotted as a dash-dot horizontal line, and for comparison the linear theory's predicted value of $2.82 \times 10^{-3} H\Omega^2$ is plotted as a dotted horizontal line. 177
- A.4 Perturbation of surface density in units of $\gamma(\gamma - 1)L/\chi c_s^2$ due to the luminous body's heat at $t = 3.5$ orbits. Contour levels on the left are a geometric series with a ratio of $\sqrt{2}$ from -0.03 to -0.48 . On the right, contour levels have a ratio of 2 between them and run from ± 0.01 to ± 0.16 . Solid contour are positive values; dashed are negative. Thermal conductivity is $\chi = 0.017 H^2\Omega$, $\ell_v = 1.42P_0\Omega$. Similar to Ref. [187, Fig. 1]. 178
- A.5 Net azimuthal acceleration due to the gas's gravity as a function of total injected luminosity. Thermal conductivity is fixed at $\chi = 0.017 H^2\Omega$. The linear theory's prediction (slope: 1.0; red dotted line) matches the data well even into the non-linear regime, although the fit (slope: $1.0 \pm 6.5 \times 10^{-7}$; black dashed line) was determined using only the linear data points. Simulations are summarized in Table A.1. 178
- A.6 Net azimuthal acceleration due to the gas's gravity as a function of thermal conductivity. Total luminosity is fixed with $L = 1.96 \times 10^{-4} P_0\Omega H^3$ ($\ell_v = 1.42 P_0\Omega$). The linear theory (red dotted line) predicts a power-law index of -1.5 , whereas the fit (black dashed line) determines a power-law index of $-1.0 \pm 1.6 \times 10^{-4}$ and total luminosity $L = 6.7 \times 10^{-3} P_0\Omega H^3$. The fit was determined solely from the simulations satisfying the hierarchies $x_p/\lambda_c < 0.3$ and $\lambda_c/H < 0.6$ (blue crosses). Simulations are summarized in Table A.2. 181

- A.7 Time-averaged azimuthal profiles at the body's position of the density (dashed line) and temperature (solid line) perturbations ρ' and T' for the fiducial simulation L1K1 (blue crosses), and a high conductivity simulation L1K10 (orange dots), and a high luminosity simulation L100K1 (green squares). Each line has been normalized to its maximum value (0.14 and 0.15 times L1K1's density and temperature perturbation maxima for L1K10; 81.3 and 137.6 times L1K1's density and temperature perturbation maxima for L100K1) and time-averaged over the last seven orbits, $t = 3$ to 9.6 orbits. Filled-in portions denote one standard deviation over time. 184
- A.8 Time-averaged azimuthal profiles at the body's position of the divergence of the heat flux's non-linear contribution $\nabla \cdot [\chi \rho' \nabla T']$ for the fiducial simulation L1K1 ($\ell_v = 1.42 P_0 \Omega$, $\chi = 0.017 H^2 \Omega$; blue crosses), the high conductivity simulation L1K10 ($\ell_v = 1.42 P_0 \Omega$, $\chi = 0.17 H^2 \Omega$; orange circles), and the high luminosity simulation L100K1 ($\ell_v = 142 P_0 \Omega$, $\chi = 0.017 H^2 \Omega$; green squares). Each line has been normalized to its maximum value (0.004 and 2.9×10^4 times the fiducial simulation's maximum value, respectively) and time-averaged over $t = 3$ to 9.6 orbits. Error bars denote one standard deviation over time. 187
- A.9 Snapshot of the radial profile of the net azimuthal acceleration on the body at a time of 9 orbits for a high conductivity simulation L1K10 ($\ell_v = 1.42 P_0 \Omega$, $\chi = 0.17 H^2 \Omega$; orange dash-dot line), and a high luminosity simulation L100K1 ($\ell_v = 142 P_0 \Omega$, $\chi = 0.017 H^2 \Omega$; green dotted line), normalized at every point to the fiducial simulation L1K1. As in Fig. A.3, the injection radius $r_{\text{rad}} = 0.04 H$ for all three simulations (black vertical line) and half the characteristic wavelength for L1K1 and L100K1 ($\lambda_c/2 = 0.26 H$) and L1K10 ($\lambda_c/2 = 0.82 H$) are plotted. Horizontal lines are the profiles that L1K10 and L100K1 would have if they had the same shape as L1K1, i.e. the linear theory's prediction. 188

- B.1 Black hole surfaces as a function of black hole spin a . The hatched area with $r < r_{\text{ISCO}}$ is the plunging region (only unstable orbits). The yellow-filled region with $r_{mb} < r < r_{\text{ISCO}}$ hosts unstable circular orbits, while the green-filled region with $r_{ph} < r < r_{mb}$ has unstable bound orbits. At $r < r_{ph}$, no circular orbits are possible, even for massless particles such as photons. The red region shows where there are stable circular orbits within the ergosphere ($r_{\text{ISCO}} < r_E$), which only occurs for large spins $a \gtrsim 0.94$ 199
- B.2 Lapse function (Eq. B.12) over radius for two different black hole spins a 201
- B.3 a) Lorentz factor γ for the Novikov-Thorne relativistic thin disk as a function of radius for two dimensionless spins. b) Ratio of fluid proper time to observer at infinity for gas with angular velocity given by the Novikov-Thorne thin disk model. 203

Chapter 1

Context and Outline

Black holes and their surroundings constitute a natural laboratory for exploring plasmas under extreme conditions that cannot yet be created on Earth. In the plasma around a black hole, conditions can range from an accretion disk cold and dense enough to thermalize both electrons and protons, to a corona hot and sparse enough that particles never interact with each other before leaving the system. While single-temperature magnetohydrodynamics (MHD) well describes the former scenario, it fails completely for the latter, where a kinetic description is instead more appropriate. Capturing both regimes in a single model requires bridging a massive separation in length scales. This scale separation spans tens of orders of magnitude, from microphysical length scales such as particle Larmor radius to length scales over which general relativity becomes non-negligible. General relativity and gas dynamics determine large-scale structure and transport magnetic fields. Microphysical processes determine the energy released by magnetic reconnection, or the fraction of dissipated energy that heats electrons rather than protons. The plasma physics problems in black hole accretion physics center around including microphysics on the scales of the accretion disk-corona system.

Understanding the physics of the plasmas around black holes is crucial to interpreting the polarization and highly resolved energy spectra from recent and upcoming X-ray telescopes such as IXPE and XRISM [1, 2]. Astrophysicists often rely on interpreting electromagnetic radiation from the gas and plasma surrounding a black hole to extract black hole properties such as mass and spin. The mass and spins of these black holes inform galactic formation and evolution models, as well as stellar evolution models. However, measuring the black hole spin currently poses a particular difficulty. Different gas-physics-dependent techniques obtain non-

overlapping estimates for the spin of accreting stellar-mass black holes. Comparing those techniques to LIGO measurements suggest that merging black holes and single black holes have different spin distributions [3]. In this sense, black hole accretion disks pose plasma physics problems with broad astrophysical implications.

This thesis explores how refining the plasma physics in accretion disk and corona models can improve our understanding of black holes and their environments. Chapter 2 introduces background information on the plasma physics and astrophysics needed to understand this work. Chapter 3 examines particle acceleration and a possible wind-launching mechanism in the collisionless regime of the corona [4]. Chapter 4 models the corona as the region closest to the black hole, crossing between the strongly collisional regime to the collisionless regime where particles can accelerate to nonthermal energies [5]. Chapter 5 performs well-controlled experiments to determine the role of Coulomb collisions between protons and electrons in determining the accretion disk and corona structure [6]. For the sake of brevity and cohesion, I include my first paper of graduate school in Appendix A [7]; although it addresses accreting bodies in a disk, i.e. stellar-mass black holes in the accretion disk around a supermassive black hole, it does not touch on aspects of the corona that I focus on in this thesis. The main scientific chapters of this thesis include a prologue where I connect the publication to the broad themes in this chapter and the relevant background information in the next chapter.

Chapter 2

Introduction and Background

In this chapter, I will review the broad physics and astrophysics relevant to the rest of the thesis. In Ch. 2.1, I will treat both collisionless plasmas and collisional plasmas. I introduce the kinetic equations, the MHD equations, and explore assumptions in plasma turbulence theory. In Ch. 2.2, I introduce the problem of accretion and the general framework for discussing black hole accretion. Ch. 2.3 delves into specific models for the accretion disk and corona combined system, exposing past assumptions that I will relax. Ch. 2.4 discusses the observational context for this thesis and the specific astrophysical objects to which this research may be applied. Finally, Ch. 2.5 outlines the numerical codes that are often crucial to studying these systems and that I employ in subsequent chapters.

2.1 Plasma Regimes

The universe hosts many types of plasmas, each with a certain ordering of length scales and dimensionless parameters that determines the mathematical formalism used to describe them. Even the narrow subset of plasmas around black holes spans several orders of magnitude in temperature and number density. Fig. 2.1 demonstrates a sampling of the parameter space of astrophysical and earthly plasmas in terms of number density and temperature. Several factors are not included in the figure for simplicity, including magnetization of the plasma^a and the entire strongly-coupled plasma regime, applicable to ultracold plasmas and the interior of gas giants [8].

On the surface, mathematically describing a plasma is simple. Plasma comprises charged particles (in

^a A magnetized plasma has a Larmor radius much smaller than the system size

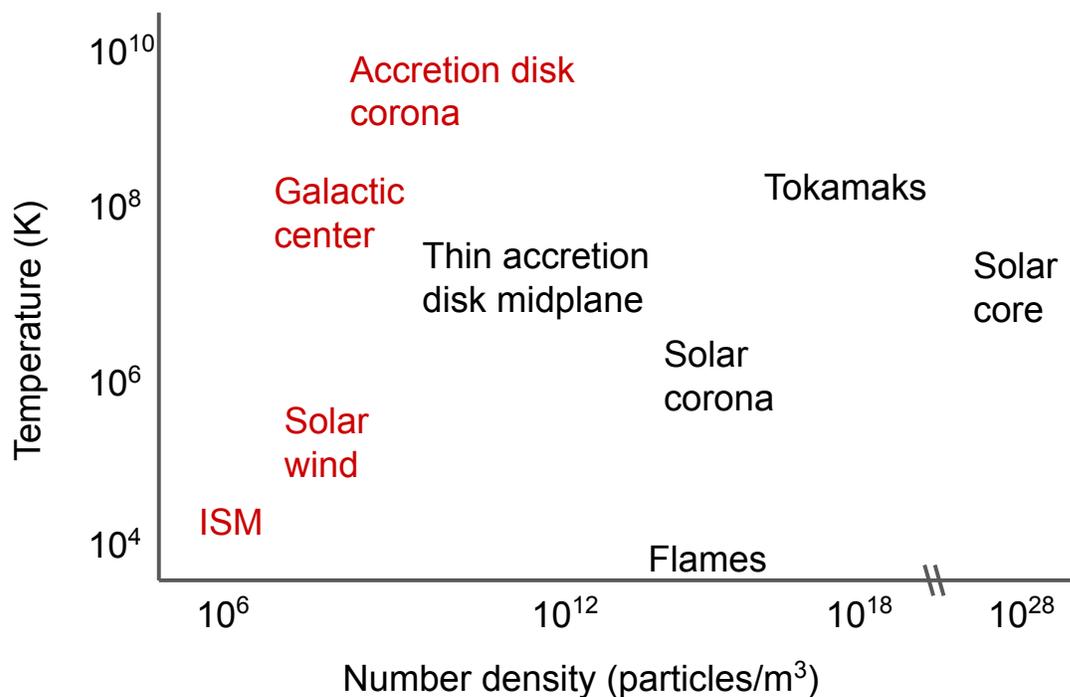


Figure 2.1: Sampling of plasma systems in number density and temperature parameter space. Astrophysical plasmas (stars excepted) generally have much lower densities than naturally-occurring plasmas on Earth, which in turn are generally sparser than laboratory-created plasmas. Text color indicates the collisionality of the system, which depends on system size in addition to density and temperature: red for collisionless and black for collisional. System parameters were estimated from Refs. [8–11].

this work, electrons and protons or electrons and positrons) that move throughout space generating electric and magnetic fields that evolve according to Maxwell's equations and affect particles through the Lorentz force. The Klimontovich equation describes such a scenario for particle density $N_s(\mathbf{x}, \mathbf{v}, t)$ with species $s \in e, i$ for electrons and protons/positrons:

$$\partial_t N_s(\mathbf{x}, \mathbf{v}, t) + \mathbf{v} \cdot \nabla_x N_s + \frac{q_s}{m_s} \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \cdot \nabla_v N_s = 0, \quad (2.1)$$

where q_s and m_s are the charge and mass of the particles for species s , ∇_x and ∇_v are the gradient with respect to position and velocity, and the electric and magnetic fields include external fields and the fields generated by the moving charges themselves [12].

In practice, Eq. 2.1 is essentially impossible to study, either analytically or numerically, for large numbers of particles. Besides, knowing how many particles are in a volume of phase space is much more useful than knowing whether a particle sits at the exact position (\mathbf{x}, \mathbf{v}) in phase space. As such, the smoother distribution function $f_s(\mathbf{x}, \mathbf{v}, t) \equiv \langle N_s(\mathbf{x}, \mathbf{v}, t) \rangle$ replaces N_s , where the brackets denote an ensemble average. By the ergodic hypothesis, the ensemble average is equivalent to a spatial average over a length scale ℓ that is much greater than the particle separation and much less than the Debye length [11]. Averaging over Eq. 2.1 yields the standard Boltzmann equation:

$$\partial_t f_s(\mathbf{x}, \mathbf{v}, t) + \left[\mathbf{v} \cdot \nabla_x + \frac{q_s}{m_s} \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \cdot \nabla_v \right] f_s = \left(\frac{\partial f_s}{\partial t} \right)_{\text{coll}} \quad (2.2)$$

where the right-hand side, known as the collision operator, includes discrete particle-particle interactions. The fields must satisfy Maxwell's equations, with the charge and current density given as moments of f_s .

2.1.1 Kinetic Description for Collisionless Plasmas

For a collisionless plasma, the collision operator on the right side of Eq. 2.2 is set to zero to obtain the Vlasov equation, also called the collisionless Boltzmann equation. Setting the collision operator to zero assumes that collisions are negligible compared to collective effects. The ratio of collisional to collective effects is given by the inverse of the plasma parameter, where the plasma parameter is the number of particles inside a Debye sphere. Because the plasma parameter is often much greater than 1, this assumption is

excellent in weakly-coupled plasmas. The red text in Fig. 2.1 highlights a few collisionless astrophysical systems where the mean free path between particle collisions is much longer than the size of the system.

2.1.2 Fluid Description for Collisional Plasmas

Taking velocity moments of the Vlasov equation yields the conservation equations of MHD. To close the resulting hierarchy that depends on the next highest moment, the fluid description assumes that the particle velocity distribution is thermal and thus characterized by a single number (temperature) at every point. Collisionless systems have a particle mean free path much larger than the system size. Because collisions push particles towards thermal equilibrium, the particles in collisional systems have a Boltzmann (non-relativistic) or Maxwell-Jüttner (relativistic) distribution [11, 12]. Two-temperature MHD allows protons and electrons to have different temperatures and to exchange energy between the species via Coulomb collisions. Single-temperature MHD assumes that the density is high enough for Coulomb collisions to maintain both species at the same temperature. The black text in Fig. 2.1 highlights a few collisional systems where particles essentially thermalize due to collisions.

2.1.3 Imbalanced Turbulence: from MHD to sub-Larmor Scales

To set the scene for Ch. 3’s study of relativistic, collisionless imbalanced turbulence, I now review the somewhat more familiar regime of incompressible imbalanced MHD turbulence. I will then extend this framework to the relevant relativistic and collisionless regimes.

Incompressible balanced (magneto)hydrodynamic turbulence has a long, rich history of phenomenological models that aim to predict the power-law scaling of the kinetic and magnetic energy spectra [13–15], the measurement of which in numerical simulations is still under debate [16–18]. By relaxing the assumption of equal energy fluxes parallel and anti-parallel to the background magnetic field — a scenario common in astrophysical systems — the turbulence becomes “imbalanced”. Breaking the symmetry along the magnetic field allows a choice in defining different quantities. Numerical models of imbalanced MHD turbulence have run into subtleties such as domain size, resolution, and long integration times [19, 20]. Due to these analytic and numerical difficulties, there is no overarching framework for understanding incompressible imbalanced

MHD turbulence.

Phenomenological models of imbalanced turbulence differ in assumptions of which length scales are relevant to the nonlinear interactions. Ref. [21], for instance, assumes that the more energetic wave's shearing rate determines the parallel length for both fields, leading to a prediction of the same anisotropy for both fields and an energy spectral index of $-5/3$. Numerical simulations suggest that this model is adequate for small levels of imbalance, but does not predict the energy spectra at larger imbalance [22]. Alternatively, Ref. [23] formulates an advection-diffusion model for the power spectrum and argues that the coherence time for the more energetic wave is on the order of $1/(k_{\parallel} v_A)$, much longer than Ref. [21]'s timescale $1/(k_{\perp} w_{-})$. As a result, this model predicts that the energy spectra have different power-law indices and that the more energetic wave's spectrum steepens with increasing cross-helicity. Ref. [24] also assumes that the more energetic wave undergoes a weak cascade; however, these last two models differ in their assumptions of the relevant length scales for the nonlinear interactions of the more energetic wave. This model appears to be consistent with numerical simulations in both the non-relativistic [22, 25, 26] and the relativistic regime [27], although the numerical methods and insufficient resolution of these simulations might have led to artificial results [28]. Lastly, Ref. [29] assumes that dynamic alignment will lead to the same nonlinear timescale for both waves, leading to power-law indices of $k_{\perp}^{-3/2}$. When simulations have a sufficient box size, resolution, and techniques to capture the inertial range, it is this model that appears to be supported numerically [20, 28, 30]. Hybrid simulations of imbalanced turbulence have focused on decaying turbulence, rather than driven turbulence [31].

More recent attempts to study imbalanced turbulence below the Larmor scale have focused on electron MHD or on kinetic Alfvén waves. Electron MHD studies have found an inverse cascade of magnetic helicity and a magnetic energy power-law index of close to $-3/2$ [32], supporting the Ref. [29] model. Studies of the transition between MHD and kinetic turbulence have shown that kinetic Alfvén wave turbulence is less imbalanced than its MHD counterpart [33]. More recently, progress in modeling collisionless Alfvénic turbulence via a diffusion equation in Fourier space suggests that high imbalance combined with Landau damping of ions leads to the steepening of the Elsässer variables' energy spectrum from $-5/3$ on scales above the sonic ion Larmor radius $\rho_s^{-1} = \sqrt{m_i/T_e}\Omega_i$ to $-7/3$ below it [34, 35]. It is still unclear how kinetic effects

will influence ultra-relativistic imbalanced turbulence.

Meanwhile, observational evidence for imbalanced turbulence comes mostly from the solar wind [36, 37]. In-situ experiments have established that the MHD Alfvénic turbulence of the solar wind is strongly imbalanced in the radially outward direction [38, 39], and that this MHD turbulence could continue to kinetic scales and heat the solar corona [40]. Other possible sources of imbalanced turbulence in the solar wind are reconnection [41] and phase-mixing [42]. However, long-standing puzzles include why the cross-helicity observed in the solar wind depends on radial distance from the sun [43, 44] and why the kinetic and magnetic energy spectral power-law index depend on cross-helicity [45]. Neither of these puzzles is completely explained by any model proposed to date (although ion-scale effects are beginning to address these problems; see [33]).

Looking back on this literature review, clearly even imbalanced MHD turbulence is not completely understood. Although application of imbalanced turbulence to high-energy astrophysical systems has previously been suggested [22], detailed exploration of the relevant parameter regime has been lacking. I will address this gap in Ch. 3.

2.2 Accretion Physics and Models

The light observed from astrophysical systems is broadly the result of accretion, wherein gravitational potential energy converts into other forms of energy, in particular electromagnetic radiation. Generally speaking, accretion in astrophysics refers to the general process of matter falling onto a central massive object due to gravity. Although accretion underlies a variety of processes ranging from galaxy formation, planet formation, and supernovae, black hole accretion is unique because black holes have no hard surface where matter can accumulate. Instead, the black hole itself is a sink of energy.

To understand the problem of accretion, consider a gas cloud at radius r_0 far from a black hole. In all likelihood, the cloud initially has some net angular momentum ℓ_0 about the black hole. Relative to the angular momentum ℓ of a circular orbit at radius r , initially the ratio $\ell_0/\ell \ll 1$, but becomes $\gtrsim 1$ at smaller r because ℓ_0 is conserved and $\ell \sim r^{1/2}$. To fall onto the black hole, the gas must lose a large amount of angular momentum. Accretion physics and models of accretion disks seek to address where and how energy

dissipates and how angular momentum is transferred away from some gas, allowing it to fall onto the central object.

Where and how energy is dissipated depends on the plasma conditions and strongly influences the radiation observed from an accreting system. Different disk regimes depend on gas density and how efficiently angular momentum is removed. These disk models span the regimes discussed earlier in Ch. 2.1 and could explain some of the observations outlined in Ch. 2.4.

2.2.1 Single-temperature, Thin Accretion Disks

The thin disk (or “alpha-disk”) model describes an accretion disk with protons and electrons in local thermodynamic equilibrium [9]. This model has high enough densities to thermalize both electrons and protons to the same temperature, and thus applies to accretion rates close to the Eddington rate^b. Crucially, the model assumes that any energy dissipated locally in the disk will immediately radiate away, keeping the plasma cool compared to the plasma of hot accretion flows (Ch. 2.2.2). Because none of the dissipated energy stays in the plasma as heat, the protons do not have much thermal pressure. Local hydrostatic equilibrium demands that the disk stay geometrically thin compared to a hotter disk, in the sense that the disk gas scale height H is much smaller than the radius from the black hole r : $H \ll r$. The optical depth in the thin disk model is much greater than one, resulting in thermal emission. See Appendix B for some relevant properties of the Kerr metric and the relativistic thin disk model.

Calculations of the local dissipation rate yield an accretion disk model whose peak energy E_{peak} depends on the black hole mass, the gas accretion rate, and the distance r from the black hole as [46, 47]:

$$E_{\text{peak}} \approx 2.9 \text{ keV} \left(\frac{L_d/L_{\text{Edd}}}{0.01} \right)^{1/4} \left(\frac{\eta}{0.1} \right)^{-1/4} \left(\frac{M}{10M_{\odot}} \right)^{-1/4} \left(\frac{r_{\text{in}}}{2} \right)^{-3/4}, \quad (2.3)$$

where $r_{\text{in}} = R_{\text{in}}/r_g$ is the disk inner edge normalized to the gravitational radius, $\eta \sim 0.1$ is the accretion efficiency, and L_d is the disk luminosity. Thin accretion disks tend to have luminosities $L_d \gtrsim 0.01L_{\text{Edd}}$, and

^b The Eddington luminosity $L_{\text{Edd}} \equiv 4\pi GMm_p c/\sigma_T$ is the theoretical maximum luminosity before radiation pressure due to electron scattering counters gravity. Although many factors modify the “true” maximum luminosity, the Eddington luminosity is a useful means to compare across systems with different black hole mass M .

can stay thin up to at least $0.3L_{\text{Edd}}$.

The mass dependence in Eq. 2.3 explains why accretion disks around supermassive black holes peak in the UV (1 – 100eV), whereas disks around stellar mass black holes peak in the soft X-rays (1-10keV). The thin accretion disk's multi-temperature blackbody describes observations of Seyfert galaxies and BHBs in the thermal soft state (Ch. 2.4.1; [48]). As will be discussed in Ch. 2.3, the thin disk could underlie the BHB hard state spectrum as well.

The canonical thin disk model assumes that the peak temperature comes from just outside the inner edge of the accretion disk, located at a slightly larger radius than the innermost stable circular orbit (ISCO; [9, 49]). The boundary condition for the ISCO states that there should be zero torque at the inner disk edge. This assumption, originally motivated by arguments of causality once the gas starts plunging towards the black hole, does not hold in the presence of large-scale torques, due to e.g. large-scale magnetic fields. Simulations questioned the validity of this assumption, ultimately inspiring new models that do include torque at or within the ISCO [50, 51].

In terms of timescales, the thin disk assumes

$$t_{\text{gas}}^{\text{cool}} = t_e^{\text{cool}} = t_{\text{heat}} \ll t_{\text{infall}}, \quad (2.4)$$

where $t_{\text{gas}}^{\text{cool}}$ is the timescale for gas (proton) cooling via Coulomb collisions with electrons, t_e^{cool} is the timescale for electron cooling due to radiation, t_{heat} is the timescale for viscous energy dissipation in the disk, and t_{infall} is the gas infall or accretion timescale. In Ch. 4, I explore how the ordering of these timescales changes within the ISCO and how the resulting spectrum can include nonthermal emission from high-energy particles. In Ch. 5, I explore how high densities can collapse a thick disk down to a thin disk, and how the resulting disk structure differs from the canonical ordering in Eq. 2.4.

2.2.2 Two-temperature Accretion Disks

At lower accretion rates and thus lower gas densities, protons and electrons can decouple and a two-temperature accretion flow develops. While the first such two-temperature model was thermally unstable [52], subsequent models introduced the idea that energy could be advected to other radii in the disk, thus stabilizing

the disk [53]. These disks should be supported by ion pressure [54], and as such are geometrically thicker than the thin disks of Ch. 2.2.1: more like a doughnut than a CD^c. In these disks, the radial pressure gradient is too important to decouple radial and vertical disk structure, which is part of why in Ch. 5 I use global simulations rather than local shearing box simulations. While the term “advection-dominated accretion flow” (ADAF) does not necessarily mean the disk is hot (slim disks are cold and advection-dominated [55]), the term ADAF usually refers to the hot flow introduced by Ref. [56], among others.

Overall, because radiative processes decrease in efficiency with smaller densities (especially two-body processes such as bremsstrahlung) and dissipated energy goes into heat instead of radiation, these disks are much less luminous than thin disks. Therefore, their radiative efficiency η can be much less than the canonical Novikov-Thorne [49] value for a given spin, meaning that for an accretion rate \dot{M} , their luminosity L will be lower:

$$L_{\text{Edd}} = \eta \dot{M}_{\text{Edd}} c^2. \quad (2.5)$$

Thick accretion flows generally describe accretion flows with accretion rates $L \lesssim 10^{-3} L_{\text{Edd}}$, but can have luminosities up to $0.3 L_{\text{Edd}}$, overlapping with the thin disk luminosity regime.

The emission from geometrically thick disks is typically optically thin, meaning that traces of the radiative mechanism are imprinted into the spectrum and polarization data. The radiative processes can in theory be traced through polarization measurements, since synchrotron emission has a polarization fraction of $\gtrsim 70\%$, although even one scattering event significantly lowers that fraction [57]. Thick disks are thought to describe low luminosity AGN such as Sagittarius A* and possibly the hard state of BHBs. Repeated Compton scatterings of synchrotron emission can sum together to yield a power-law with photon index 1.7 - 1.9, in agreement with the hard state [58]. Nonthermal particles could also exist in such two-temperature flows, both within the disk [59] and in the upper, sparser layers.

Canonical ADAF models often ignore the difference in cooling times between electrons and protons. Protons can only cool through Coulomb collisions with electrons because of their larger mass, whereas electrons cool via synchrotron, bremsstrahlung, and inverse Compton radiative processes. The hierarchy of

^c As of 2023, the author does still own a CD player.

timescales reads:

$$t_{\text{infall}} \ll t_{\text{gas}}^{\text{cool}} \sim t_e^{\text{cool}} \sim t_{\text{heat}}. \quad (2.6)$$

Compare to the thin disk ordering of Eq. 2.4. The disks are radiatively inefficient because $t_e^{\text{cool}} > t_{\text{infall}}$ and advection dominated for $t_{\text{gas}}^{\text{cool}} > t_{\text{infall}}$. In Ch. 5, I will explore the possibility of a new ordering of timescales that should be relevant for the intermediate mass accretion regime. Namely, how does the disk structure change when $t_{\text{gas}}^{\text{cool}} > t_{\text{infall}} > t_e^{\text{cool}}$?

2.2.3 Magnetic Fields in Accretion Disks

Magnetic fields play a crucial role in accretion physics. Early estimates demonstrated that molecular friction between gas rings at different radii is orders of magnitude too small to explain the luminosity observed from accreting systems [46]. As such, other mechanisms (most often involving magnetic fields in some way) must make up the majority of the angular momentum loss. One popular mechanism is the magnetorotational instability, wherein magnetic fields couple plasma at different radii and the magnetic tension acts as a spring to transfer angular momentum [60, 61]. Magnetic fields can also drive winds off the accretion disk that carry away angular momentum. A common wind-launching mechanism, the magnetocentrifugal mechanism, treats the plasma as a bead on a wire, where the wire is a large-scale magnetic field [62]. In GR simulations of the accretion disk, the highly-magnetized region above the accretion disk can account for a significant fraction of angular momentum removal [63–65]. The presence of magnetic fields in accretion disks opens up the possibility of MHD turbulence, magnetic reconnection, and particle acceleration. In particular, magnetic fields can affect the distribution of dissipated energy between electrons and protons [66–68], with significant observational consequences [69, 70]. In Ch. 3, I explore how magnetic turbulence in the presence of asymmetric energy injection, like the case above a thin accretion disk, can affect the turbulent cascade and particle acceleration.

2.3 Disk-Corona Models

In a global sense, disk-corona models should explain the energy budget of the corona, in particular how it is heated after losing energy to radiation, and how/whether it contributes to angular momentum loss,

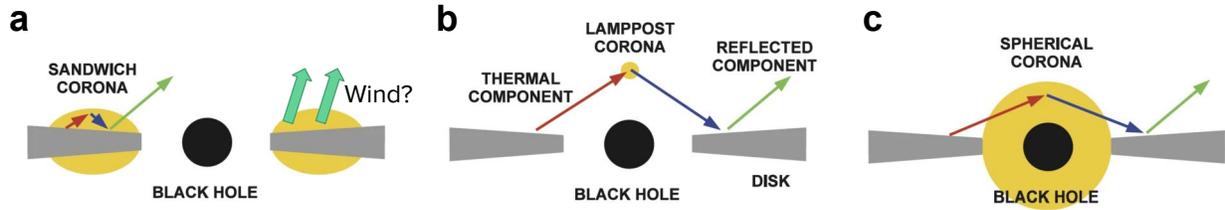


Figure 2.2: Possible configurations of the disk/corona system. The thermal disk emission (red) is upscattered by hot electrons in the corona. The resulting high-energy photons (blue) reflect off the disk, which then emits “reprocessed” light (green). Adapted from Ref. [47].

possibly through a disk wind. Disk-corona models need to address both the radiation mechanisms that lead to the observed optically thick and optically thin emission outlined in Ch. 2.4 and fit into or build upon existing accretion disk models (Ch. 2.2.1 and 2.2.2).

2.3.1 Geometric Configurations

The geometry of the corona remains largely unknown, with several configurations that match the observational suggestion of a compact corona (Ch. 2.4.2). The main questions center around whether the corona is outflowing (i.e. a jet or wind) and whether it is located above the disk (i.e. a lamppost/sandwich model) or is a truncation of the accretion disk (i.e. a hot accretion flow within a thin accretion disk that ends at radius $r > r_{\text{ISCO}}$). Fig. 2.2 illustrates some of the possible geometries that I will discuss below.

Numerical simulations of thin accretion disks often find the formation of a strongly magnetized region above the main disk body, similar to a “sandwich” type model (Fig. 2.2a). This sandwich corona can be in hydrostatic equilibrium or outflowing. The thin accretion disk at the core might not even be necessary to reproduce the hard spectral state; a hot accretion flow pushed to intermediate mass accretion rates that produce optical depths close to 1 and large Compton- y parameters can produce the power-law through repeated Compton scattering [58]. Outflowing models of the corona suggest that the power-law spectral shape originates from Comptonization off a relativistic bulk flow. For example, the lamppost model (Fig. 2.2b) suggests that the corona forms the base of a jet [71] and is widely used because of its simplicity and the ease of measuring a “height” of the corona. Some observations that rely on lamppost models suggest

that the corona expands away from the black hole during a BHB state transition [72]. Another model suggests that reconnection in the magnetically-dominated region both heats the corona [73, 74] and launches plasmoids that form the relativistic outflow [75].

Disk truncation models propose that the corona comprises a hot, thick accretion flow surrounded by a cold, thin accretion disk that truncates at some radius that is typically greater than the ISCO (Fig. 2.2c). These models are attractive for explaining BHB state transitions (Ch. 2.4.1), the idea being that the thin accretion disk moves in or out as the system transitions from quiescent to hard and soft states [76]. Fig. 2.3a illustrates the disk truncation model as a function of mass accretion rate. Although the corona is often thought to disappear in the soft state, it could just shrink to a small size and low luminosity, as suggested by some observations [77] and explored in Ch. 4. See Ref. [78] for a review of observational evidence for the disk truncation model. Disk truncation has been numerically reproduced a handful of times, for example by introducing an artificial cooling function to induce it at a given location [79], evolving a highly-magnetized disk [80], or introducing cooling into an alpha-disk model [81]. The truncation mechanism in these simulations differs between thermal instability and magnetic torques. In Ch. 5, I will explore the possibility that Coulomb collisions becoming inefficient determines the truncation radius.

2.3.2 Plasma Physics in the Corona

Regardless of the geometric configuration, the hot corona probably involves a collisionless, strongly magnetized plasma that suggests the presence of many interesting processes. These processes fall broadly into two categories: those that *heat* the plasma, i.e. retain its thermal distribution but increase its temperature, and those that *accelerate* particles, changing the shape of the particle distribution. Fig. 2.4 illustrates the distinction between these processes for an initially thermal particle distribution. The occurrence and dominance of one process over another are broadly determined by the plasma's physical conditions and the sources and sinks of energy in the system. Potential sources of energy include turbulence injected from a thin disk (see Fig. 2.3b) and magnetic fields, which can lead to conversion of magnetic energy to kinetic energy via mechanisms like turbulence and magnetic reconnection. The dominant energy sink is likely radiation, though other sources such as conduction with the disk [82] are in principle possible. Of course, because

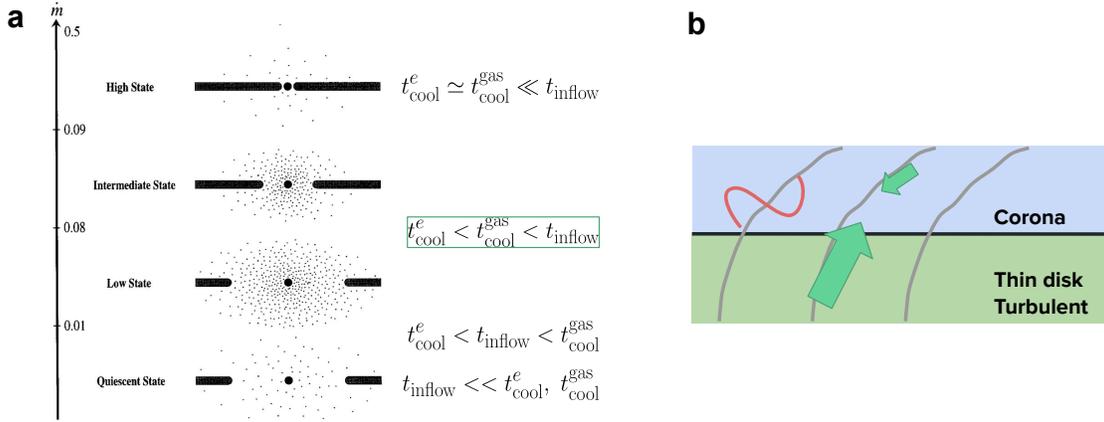


Figure 2.3: a) An illustration of the disk truncation model for BHB state transitions. At high luminosities and accretion rates, a thin accretion disk reaches all the way to the ISCO. At smaller luminosities, the thin disk truncates and a thick disk populates the inner region. Ch. 5 investigates the intermediate regime. Figure adapted from Ref. [76]. b) Schematic of a possible mechanism for particle acceleration in a sandwich-type corona (blue) sitting above a thin disk (green) threaded by large-scale magnetic fields (gray). Alfvén waves (red) launch predominantly upward into the corona and then reflect. Green arrows show the dominant direction of energy flux that starts the turbulent cascade. Ch. 3 investigates heating and particle acceleration in this scenario.

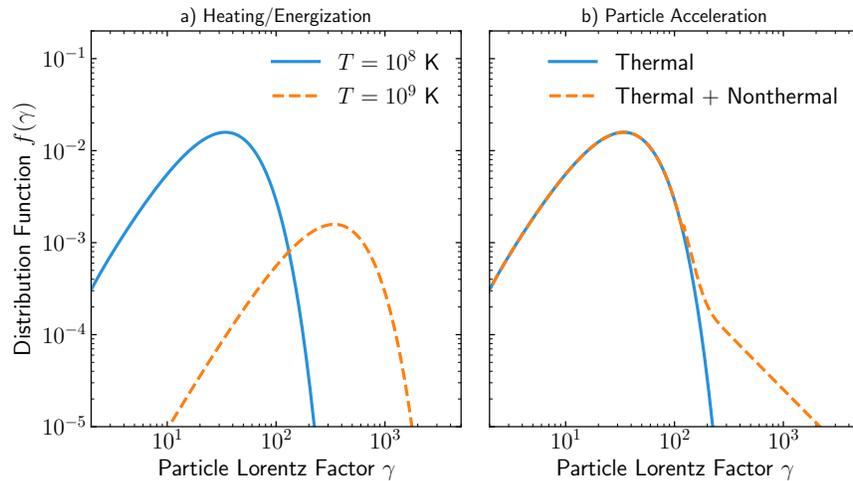


Figure 2.4: Illustration of a) plasma heating/energization and b) nonthermal particle acceleration. Energy injected into a system with an initial distribution function in blue can end with a final distribution function shown in orange. Electrons and protons do not need to experience the same effect from each process. Here I assume an isotropic particle distribution function and have normalized to the particle number density.

current experiments cannot make *in-situ* measurements of the particle distribution, we must be careful relying on observed radiation to disentangle the underlying particle distribution.

The heating problem of accretion disk coronae resembles the problem with the solar corona: farther away from the core of the sun, the temperature actually increases, contrary to naive expectations. Similarly, for a sandwich-type corona, the temperature increases with distance from the disk midplane by about two orders of magnitude. The heating mechanism for both accretion disk coronae and the solar corona could lie in Alfvénic dissipation [83, 84], wherein Alfvén waves launch from the disk/sun into the corona, reflect off density gradients or perturbations, and subsequently form a turbulent cascade — see the schematic in Fig. 2.3b. In addition to turbulence, magnetic reconnection could also heat the plasma. Both turbulence and magnetic reconnection could also accelerate particles to high energies, resulting in nonthermal emission.

In the relativistically-hot, magnetically-dominated regime, magnetic turbulence and reconnection can efficiently accelerate particles [68, 85, 86]. Although the possibility of nonthermal electrons in the corona was long-established [87], full numerical models of the accretion disk-corona system with nonthermal electrons are rare (see Ch. 2.5, [88]). The fraction f_e of dissipated energy that heats and accelerates electrons so far varies depending on factors such as plasma magnetization [66, 68, 89], magnetic field configuration and in particular the guide field [90], and the turbulence driving mechanism [91, 92]. The fraction f_e has long been a subject of debate in the accretion disk community. It could affect the radiation observed from the disk and jet [58, 69]. If the temperature of the protons is significantly affected by f_e , then the structure of the disk and the shape of the corona could also change depending on this fraction. Particle-in-cell simulations (PIC; Ch. 2.5.2) have started to address the problem of f_e 's values as a function of relevant parameters such as plasma magnetization and magnetic field configuration.

2.4 Observational Motivation: Thermal and Nonthermal Emission

In this thesis, I will seek to understand processes occurring light years away, around black holes far beyond our own solar system. Most black holes and their accretion disks are too far away to spatially resolve, leaving theorists to interpret spectra and occasionally polarization data. Astronomers often split spectra observed from the vicinity of a black hole into two components: a thermal distribution and a nonthermal,

often power-law distribution. This separation is motivated by the idea that the thermal component comes from a thin accretion disk (see Ch. 2.2.1) and the power-law component comes from a “corona” — some sort of hot gas with an unknown geometry. Although this separation glosses over the source of the corona and whether there actually is an underlying thin accretion disk, the general concept of optically thick (thermal) emission and optically thin emission (power law) applies.

While both supermassive black holes and stellar-mass black holes host accretion disks and coronae, the shorter evolution timescales for stellar-mass black holes mean we can see more time-varying behavior for them. As such, although much of my work applies to the disks in Seyfert galaxies, I will primarily focus on black hole binary (BHB) systems.

2.4.1 Black Hole Binary State Transitions

In a black hole binary, a star accretes onto a stellar-mass black hole either through Roche lobe overflow or stellar winds. The star will emit in the IR/optical/UV with a typical blackbody spectrum. The X-ray emission, on the other hand, undergoes dramatic outbursts that completely alter the bolometric luminosity and spectral shape. At first, the system is in quiescence with a low luminosity $L \ll L_{\text{Edd}}$. On the first leg of the outburst, the X-ray luminosity increases dramatically (Fig. 2.5a). Over the course of days to weeks, the spectrum morphs from peaking at 100 keV (Fig. 2.5b blue) to a blackbody with a temperature of 1 keV (Fig. 2.5b red), then back, eventually returning to quiescence. This cycle repeats approximately every 1 to 60 years depending on the system and shows hysteresis in the return to quiescence; the transition from the 100 keV to the 1 keV peaked states happens at a different luminosity than the transition from the 1 keV to the 100 keV state [48].

To facilitate categorization of these outbursts into distinct spectral states, astronomers rely on the division of the spectrum into thermal and nonthermal (power-law) components. The thermal distribution yields a characteristic temperature, while the power-law photon distribution $N(E) = A_E E^{-\Gamma}$ fits the normalization A_E and the *photon* power-law index Γ . The temperature is often used to make arguments about how close to the black hole the thin accretion disk potentially extends, while the power-law index could provide information about the underlying radiation processes (synchrotron, inverse Compton) and gas densities. The

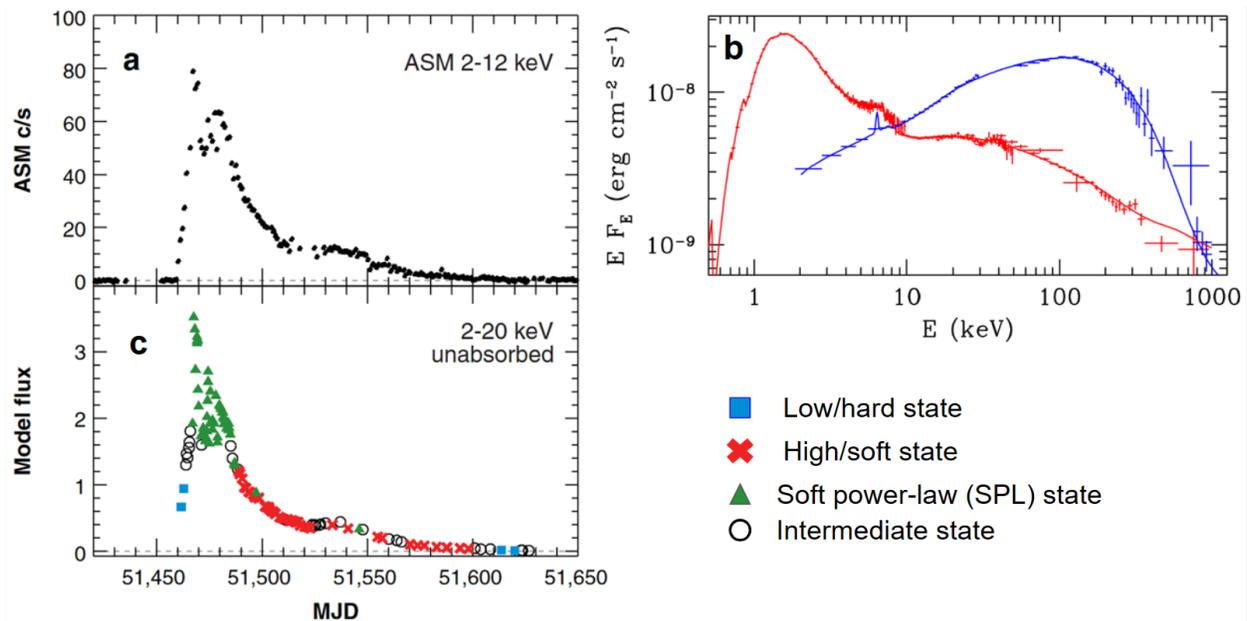


Figure 2.5: Black hole binaries undergo outbursts that significantly change their X-ray luminosity and spectral shape. Panel a shows the counts per second over time (Modified Julian Date; MJD) detected from the binary XTE J1859+226 by the All-sky Monitor (ASM) aboard the Rossi X-ray Timing Explorer (RXTE), which dramatically increase over the course of an outburst. Panel b shows spectra recorded from observations of the binary Cyg X-1 during two distinct X-ray states: the high/soft state in red and the low/hard state in blue. Panel c demonstrates how model fits to spectral data from panel a characterize the time evolution of the binary's spectral states. Adapted from Ref. [48] and Ref. [78].

State Name	Definition
Thermal or High/soft	Disk fraction $>75\%$
Low/hard	Disk fraction $<20\%$ $1.4 < \Gamma < 2.1$
Steep Power Law (SPL) or Very High State (VHS)	Disk fraction $\lesssim 0.8$ $\Gamma > 2.4$

Table 2.1: Definitions of black hole binary spectral states. Adapted from Ref. [48]

ratio between the thermal and power-law flux (often called the “disk fraction”) and the power-law index Γ together empirically define the spectral states (see Table 2.1). Measuring these quantities in different energy bands and taking the ratio also yields a hardness ratio, a ratio between more energetic (“hard”) and less energetic (“soft”) X-rays^d that is used to track the state transitions.

As discussed in Ch. 2.2, a changing accretion disk well describes these outbursts, although the details of the emission and the mechanism behind the transition remain less understood. For example, the origin of the nonthermal emission across different accretion states could vary from saturated Comptonized emission to synchrotron radiation. My research describes a variety of emission models in different accretion states, sometimes focusing purely on the power-law component and sometimes holistically describing the disk fraction by comparing dissipation in optically thick vs. optically thin regions.

2.4.2 Spectroscopy and Timing Studies of the Disk-Corona System

Spectra from black hole accretion disks often reveal features that are thought to come from light reflecting off a hot corona rather than coming directly from the accretion disk (Fig. 2.2). Hot electrons (either thermal or nonthermal) will Compton upscatter a population of seed photons (possibly thermal emission from the disk or nonthermal synchrotron photons). Some portion of these higher energy photons will then reflect off of the accretion disk, ionizing atoms like iron, whose fluorescence lines then appear in the X-ray spectrum. Because the reflection presumably happens close to the black hole, the reflection spectrum provides information about the innermost region of the accretion disk.

To demonstrate how delicate inferring black hole and accretion disk parameters from the reflection

^d The exact energy bands depend strongly on the instrument.

spectrum is, Fig. 2.6a shows a set of X-ray observations that feature these reflection lines, fit by assuming a truncated thin disk (see Ch. 2.3) and a hot reflection component. This fit finds an inner disk radius of $17 \pm 3r_g$, between 3 – 10 times the size of the ISCO, with a temperature of 0.15 ± 0.01 keV and a coronal temperature of 98 ± 7 keV [93]. The reflection spectrum blurs together once relativistic effects are included (Fig. 2.6b), opening up the possibility that the asymmetric relativistic broadening of the iron $K\alpha$ line could be used to measure black hole spin. This blurring is simultaneously useful to estimate black hole spin, and extremely uncertain due to the limited spectral resolution of current instruments. XRISM, launching in late 2023, is expected to distinguish between spectral lines [2]. In the meantime, Fig. 2.6c shows how current models rely heavily on assumptions about the accretion disk structure, motivating the research in Ch. 5.

In theory, changes in the disk emission will also affect the reflection spectrum after a time that depends on the spatial separation between the disk and the corona. Reverberation mapping tracks the temporal evolution of the disk and reflected emission to extract this distance and hence attempt to interpret the disk/corona geometry. Recent observations suggest that the corona is relatively compact, within $5r_g$ for a BHB [72] and $6r_g$ for a Seyfert galaxy [95]. This compactness will motivate Ch. 4’s model for the corona as a collisionless gas within the ISCO, and Ch. 5’s investigation of whether Coulomb decoupling in the inner accretion disk regions leads to the formation of a corona-like structure.

Although promising, reverberation mapping relies on several assumptions that may not hold in the actual disk/corona geometry. In particular, most techniques assume that the light-crossing time is the most relevant timescale in the system, i.e. that the region is extremely optically thin, and that the emission comes from a single central source [96]. However, both assumptions might fail in the case of a real corona sitting around the disk, either through an extended geometry or a moderate optical depth [58, 97].

2.5 Numerical Tools

Astrophysics and astronomy as a field has the curious problem that scientists cannot do the controlled experiments that are the bread and butter of “regular” physics. Astronomers cannot create a star in the lab, much less tweak a single variable and see how that change affects the evolution and death of the star. As astrophysicists, we have only what the universe gives us, and all the observational biases that come with

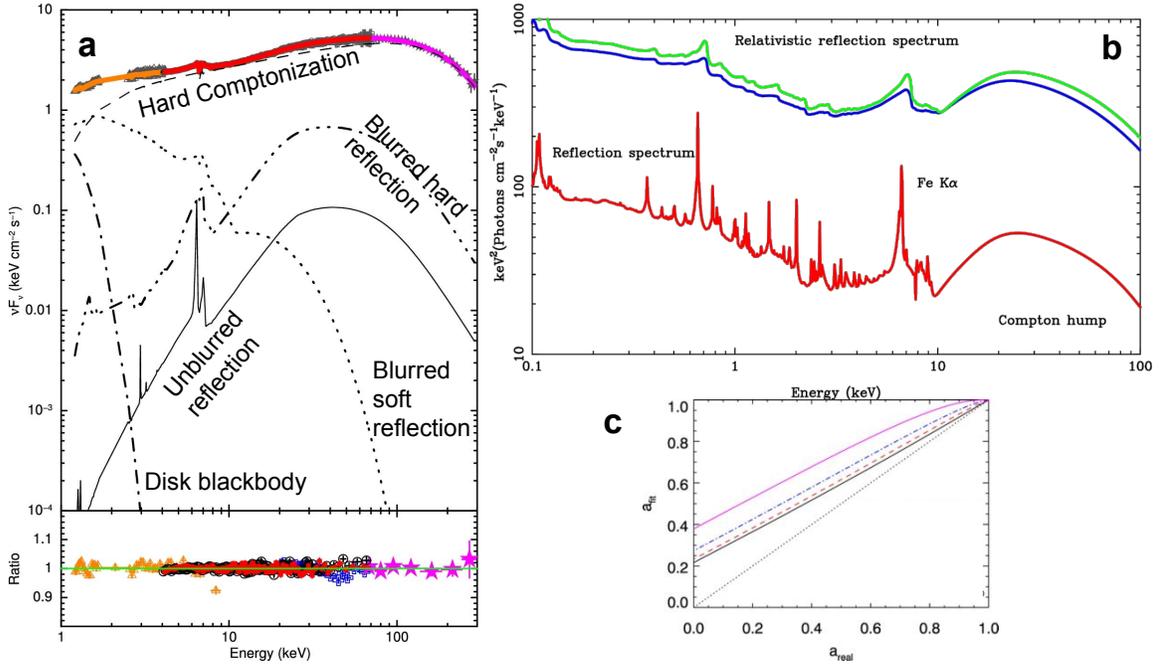


Figure 2.6: Interpreting observational data to extract black hole spin relies on fitting several components of the reflection spectrum, which is assumed to come from hot, presumably thermal electrons located somewhere around the black hole. Relativistic effects, which depend sensitively on how close to the black hole the radiation travels, blur the spectrum further and form the basis for iron-line-based spin measurements. a) Actual observations from the Suzaku, with different detectors shown in orange, red, blue, and pink. Model component fits are labeled and shown in black. The bottom panel shows the ratio of the data to the model. Adapted from Ref. [93]. b) Example reflection spectrum before (red) and after (blue/green) relativistic blurring. The green line shows the relativistically blurred spectrum for a non-spinning black hole, while the blue line shows the spectrum for a black hole spin of $a = 0.998$. From Ref. [47]. c) Current models can drastically overestimate black hole spin, motivating the need for better understanding of the plasma physics in the corona. Data from a 3D MHD simulation, scaled to have an input spin a_{real} , is fed through a model that measures spin through the fluorescent iron broadening. The model's extracted spin a_{fit} is shown on the y-axis. The dotted black line shows the perfect recovery case, while colored lines show the dependence on disk scale height at the ISCO. From bottom to top: $h/r_g \in 0.01, 0.25, 0.5, 1.0$. From Ref. [94].

trying to figure out what we are seeing. Because of lack of direct experimentation, astrophysicists often turn towards numerical codes to integrate the differential equations relevant for their system. By starting with a different initial condition, turning off a term, adding new terms, etc., theorists design and conduct numerical experiments to give us physical insight. Of course, we must be careful with numerical simulations because we cannot discover something that has not already been written into the equations we tell the computer to solve. As is the case with all simulations, different jobs need the right tools. In the case of plasma physics, the collisional, magnetohydrodynamic regime and the collisionless, kinetic regime motivate different numerical formalisms.

2.5.1 Collisional Plasma: General Relativistic Magnetohydrodynamic Codes

General Relativistic Magnetohydrodynamic (GRMHD) simulations are useful for examining the global dynamics of an accretion disk. While these simulations cannot resolve microphysics such as magnetic reconnection or accelerate particles to nonthermal energies, they can capture large-scale gradients in gas density and magnetic fields. Although MHD is strictly valid only for the collisional regime described in Ch. 2.1.2, it is much less computationally expensive than PIC and thus is commonly used to simulate an entire accretion disk.

GRMHD codes solve the general relativistic equations for conservation of particle number and energy/momentum and the evolution of the electromagnetic field tensor in the ideal limit. The first two of these equations are given by:

$$\partial_t (\sqrt{-g}\rho_0 u^t) = -\partial_i (\sqrt{-g}\rho_0 u^i) \quad (2.7)$$

$$(T^\mu_\nu + R^\mu_\nu)_{;\mu} = 0 \quad (2.8)$$

where ρ_0 is the co-moving rest mass density, u^μ is the fluid four-velocity, and R^μ_ν is the radiation stress-energy tensor. The stress-energy tensor T^μ_ν in the ideal MHD limit is written as

$$T^\mu_\nu = \left(\rho_0 + u + P + b^2\right) u^\mu u_\nu + \left(P + \frac{1}{2}b^2\right) g^\mu_\nu - b^\mu b_\nu \quad (2.9)$$

where P is the (assumed isotropic) fluid pressure, u is the fluid internal energy density, $b^2 = b^\mu b_\mu$, and b^μ is the magnetic field four-vector.

The magnetic field evolution is governed by Maxwell's equations:

$$\partial_i (\sqrt{-g} B^i) = 0 \quad (2.10)$$

$$\partial_t (\sqrt{-g} B^i) = \partial_j (\sqrt{-g} [b^j u^i - b^i u^j]) \quad (2.11)$$

where B^i is the three magnetic field, which includes the constraint $b^\mu u_\mu = 0$ and gives two equations relating b^μ and B^i : $b^t = B^i u^\mu g_{i\mu}$ and $b^i = (B^i + b^t u^i)/u^t$. See Ref. [98–100] for more details.

The treatment of the radiation term R^μ_ν , differs across GRMHD codes and often restricts the code to certain optical depth regimes. A common method for the optically thin regime uses photon packets to sample the phase space and capture emission and scattering. This Monte Carlo method quickly becomes numerically expensive at high optical depths. On the other hand, radiation can be treated as a fluid in the optically thick regime with the so-called ‘‘M1 closure’’, which closes the radiation GRMHD equations by assuming that the radiation tensor is isotropic in the orthonormal rest frame of the radiation (distinct from the fluid frame in GR) [101]. The M1 closure method does not accurately treat anisotropic radiation and as such does not apply in optically thin regions where radiation is not traveling in one main direction. Connecting the optically thin and thick regimes is thus a major numerical difficulty — and an important problem, as that is exactly the regime thought to occur during the hard-to-soft spectral state transitions described in Ch. 2.4. In Ch. 5, I will approximate radiation as a cooling function in an attempt to capture both the optically thin and optically thick regimes.

In Ch. 5, I relax the typical assumption that electrons and protons have the same temperature. To allow for electrons having a different temperature than protons, the code separately evolves an electron fluid on top of the total (electron and proton) fluid. This electron fluid is governed by the electron entropy conservation equation:

$$\frac{\rho^{\gamma_e}}{\gamma_e - 1} u^\mu \partial_\mu \kappa_e = f_e Q_H + Q_C(T_e, T_p) - u^\nu R^\mu_{\nu;\mu} \quad (2.12)$$

where $\kappa_e \equiv \exp((\gamma_e - 1)s_e) = P_e/\rho_0^{\gamma_e}$, where s_e is the electron entropy. The fraction f_e describes how much of the viscous heating Q_H goes into electrons: the remaining $(1 - f_e)Q_H$ goes into the protons. Energy is exchanged between the electrons and protons via the Coulomb heating term Q_C , which depends on the electron and proton temperature and is given by Ref. [102]. The last term on the right side is the radiation

reaction on the electron fluid. The electron fluid is described in detail in Ref. [103, 104].

2.5.2 Collisionless Plasmas: Particle-in-cell Codes

Particle-in-cell (PIC) simulations focus on the microphysics of a plasma. Instead of capturing large-scale gradients and general relativity, PIC simulations resolve the plasma frequency and do not assume a thermal particle distribution function. As such, PIC simulations best describe the collisionless regime outlined in Ch. 2.1.1. Due to their computational cost, PIC simulations often zoom in on a small patch of plasma over which the plasma and spacetime conditions do not change significantly. A notable exception to that limitation is GRPIC simulations [105–107]; note, however, that those simulations (so far) most often include an infinitely thin accretion disk.

Particle-in-cell simulations do not evolve the full distribution function $f(\mathbf{x}, \mathbf{v})$. Instead, they sample the distribution function and track macroparticles that represent many individual particles. The electromagnetic fields are solved on a grid using Maxwell’s equations. The particles are then evolved according to the Lorentz force, and their movements update the currents and charges in the next evolution of Maxwell’s equations [108, 109]. PIC simulations do not typically (or ever, in astrophysics) include particle-particle interactions such as Coulomb collisions, since such a binary interaction would significantly increase the computational cost. They occasionally include radiation, in particular synchrotron radiation as a drag force [109] and inverse Compton scattering as either a soft photon bath that produces a drag on each particle [110] or by tracking macropotons [111].

In Ch. 3, I use PIC simulations to investigate how relaxing the assumption of isotropic energy injection affects the turbulent cascade [13, 14] as described in Ch. 2.1.3.

2.6 Summary of This Work

In this thesis, I bring together both collisionless and collisional plasma physics to study the accretion disks and coronae around black holes. Kinetic physics has been especially neglected in the past, partially due to a lack of understanding and partially due to the difficulty in including the results in MHD simulations. To begin bridging these regimes, I begin with a study of the kinetic physics in the corona using PIC simulations

(Ch. 3). I examine plasma heating and particle acceleration in a corona pumped by turbulence from an underlying accretion disk. Next, I create a semi-analytic model that begins with a highly-collisional thin disk and show that kinetic physics could become important within the black hole's ISCO (Ch. 4). Particle acceleration to nonthermal energies could occur and be observed from this region, forming a version of the corona for the soft state of BHBs. Finally, I end with an exploration of two-temperature MHD in an entire accretion disk (Ch. 5). I explore the impact of Coulomb collisions between electrons and protons on determining the overall disk structure, particularly the transition between an inner hot accretion flow and an outer cold accretion disk. Throughout this thesis, I focus on the plasma physics that could affect what we observe from black hole accretion disks.

Chapter 3

Kinetic Simulations of Imbalanced Turbulence in a Relativistic Plasma: Net Flow and Particle Acceleration

Observations of astrophysical systems such as black hole binaries suggest the presence of nonthermal electrons with suprathermal energies. These high-energy electrons must somehow be accelerated out of the thermal bulk. A possible mechanism for accelerating particles is via magnetic turbulence. However, canonical models of turbulence assume that the energy flux parallel and anti-parallel to a background magnetic field is equal, or “balanced”, which might not be the case in systems such as pulsar wind nebulae and accretion disk coronae. In particular, in a sandwich-type corona, turbulence from an underlying thin disk could launch Alfvén waves along magnetic field lines. These waves could then reflect and form counterpropagating waves: see Fig. 2.3b.

Here, I investigate how imbalance in energy injected into a magnetized, relativistic, collisionless plasma can affect the turbulent cascade’s ability to heat the plasma and accelerate particles to nonthermal energies. See Ch. 2.1.3 for a review. To effectively understand these microphysical plasma processes, I employ PIC simulations (Ch. 2.5.2) to study a small patch of plasma. In these simulations, I ignore general relativistic effects and large-scale gradients in the disk. Future work will investigate the impact of large-scale gradients. I find that imbalance decreases the efficiency of particle acceleration, effectively limiting the development of the turbulent cascade. A portion of the imbalanced net injected energy converts into net motion of the plasma, which could lead to an outflow. Such an outflow could affect the structure of the corona. This chapter was published in Ref. [4].

Abstract

Turbulent high-energy astrophysical systems often feature asymmetric energy injection: for instance, Alfvén waves propagating from an accretion disk into its corona. Such systems are “imbalanced”: the energy fluxes parallel and anti-parallel to the large-scale magnetic field are unequal. In the past, numerical studies of imbalanced turbulence have focused on the magnetohydrodynamic regime. In the present study, we investigate externally-driven imbalanced turbulence in a collisionless, ultrarelativistically hot, magnetized pair plasma using three-dimensional particle-in-cell (PIC) simulations. We find that the injected electromagnetic momentum efficiently converts into plasma momentum, resulting in net motion along the background magnetic field with speeds up to a significant fraction of lightspeed. This discovery has important implications for the launching of accretion disk winds. We also find that although particle acceleration in imbalanced turbulence operates on a slower timescale than in balanced turbulence, it ultimately produces a power-law energy distribution similar to balanced turbulence. Our results have ramifications for black hole accretion disk coronae, winds, and jets.

3.1 Introduction

High-energy astrophysical systems such as accretion disks, jets, and pulsar wind nebulae often comprise collisionless, relativistically-hot plasmas and are likely turbulent [9, 61, 112, 113]. Turbulence in systems with magnetization (the ratio of magnetic enthalpy to plasma enthalpy) $\sigma \gtrsim 1$ can efficiently accelerate particles, as recently demonstrated in particle-in-cell (PIC) simulations [86, 114–116]. Such nonthermal particle acceleration (NTPA) could explain the power laws seen in spectra of jets, pulsar wind nebulae, and stellar-mass black hole X-ray binary systems [48, 117, 118].

These previous studies of turbulence in relativistic collisionless plasmas have assumed symmetric energy injection into the plasma. However, this assumption is not true in a variety of space and astrophysical systems where turbulence is preferentially stirred on one side of the system. For example, in an accretion disk-wind system, turbulence in the disk may shake the footpoint of an open large-scale magnetic field line, sending Alfvén waves predominately away from the disk’s midplane and into the corona [84]. This

asymmetric propagation of Alfvén waves could impact NTPA. In addition, if efficiently coupled to the plasma, such asymmetrically-injected electromagnetic momentum could result in bulk motion of the plasma—an outflow/wind. Understanding both NTPA and the possible formation of an outflow necessitates studying turbulence with asymmetric momentum injection—so-called “imbalanced” turbulence.

Most studies of imbalanced turbulence have focused on the magnetohydrodynamic (MHD) regime. Canonical phenomenological models for strong, “balanced” MHD turbulence consider ensembles of counter-propagating Alfvén waves with equal energy fluxes along a background magnetic field [14, 15, 119–121]; see Ref. [17] for a recent review. Phenomenological models for imbalanced turbulence relax the assumption of equal fluxes, leading to predictions that are ripe for numerical exploration [21, 23, 24, 29]. Numerical attempts to model imbalanced turbulence in the MHD regime have proven difficult due to questions about the effects of varying dissipation prescriptions, limited dynamic ranges of accessible simulation domains, and limited run times [19, 20, 25, 26, 28, 30]. Some numerical studies have extended beyond standard MHD to the relativistic MHD regime using the force-free assumption that $\sigma \gg 1$ [27, 32].

Below MHD scales, analytic models of imbalanced kinetic turbulence have recently been formulated in the nonrelativistic regime [33, 35, 122]. These models of collisionless imbalanced turbulence can be tested against measurements of the solar wind [45]. Meanwhile, numerical studies have made approximations of infinite ion-to-electron mass ratio to model scales below the proton gyroradius [123], demonstrated the importance of finite Larmor radius effects on the turbulent energy cascade [124], or employed a diffusive model to study turbulence from fluid to sub-ion scales [34]. A few numerical studies have modeled the fully kinetic collisionless regime with an eye towards the solar wind [125]. However, none to our knowledge have examined the ultrarelativistic, collisionless regime relevant to high-energy astrophysical systems. Studying this regime is important because quasi-linear models of turbulent particle acceleration for imbalanced MHD predict a decrease in the Fokker-Planck momentum diffusion coefficient for increasing imbalance, leading to less efficient NTPA and posing a potential obstacle to turbulence as an astrophysical particle accelerator in some systems [126, 127].

In this work, we explore imbalanced, relativistic turbulence in magnetized collisionless electron-positron (pair) plasmas using 3D PIC methods [128]. We study how imbalance affects self-consistent NTPA,

inaccessible in fluid-based models, and how it introduces an effect entirely absent in balanced models: the transfer of net momentum to the plasma, which in realistic systems could form outflows. Though the regime we simulate is particularly applicable to black hole accretion disk coronal heating [84] and wind-launching, our results should be generally applicable to relativistic astrophysical turbulence where the source of perturbations is localized, such as the jets originating from active galactic nuclei.

To frame our study of this numerically- and analytically-unexplored regime of turbulence, we focus on four main questions:

- (i) How does imbalance affect the formation of a turbulent cascade?
- (ii) How does imbalance affect the partition of large-scale injected energy into electromagnetic, internal, and turbulent kinetic energy?
- (iii) Does imbalance drive net motion of the plasma?
- (iv) How does imbalance affect NTPA?

We first introduce the numerical tools and parameters used to describe imbalanced turbulence (Section 3.2). We then present the results of 3D collisionless, relativistic PIC simulations with varying degrees of imbalance and ratio of system size to initial Larmor radius that address each of the above questions in order. After first demonstrating the presence of a turbulent cascade (Section 3.3.1), we discuss how and why the energy partition changes with imbalance (Section 3.3.2). Then we examine the formation of a net flow via efficient momentum transfer to the plasma and provide an analytic framework for understanding it (Section 3.3.3). We continue by demonstrating, for the first time, the similarity of NTPA in balanced and imbalanced turbulence (Section 3.3.4). We check the dependence of our results on simulation domain size in Appendix 3.A. We conclude with implications for high-energy astrophysical systems and remaining questions (Section 3.4).

3.2 Methods

In this section, we will first review physical properties of relativistic magnetized plasmas (Section 3.2.1), and then outline the simulation suite used to study imbalanced turbulence (Section 3.2.2). The last subsection discusses various diagnostics that will be used to analyze our simulations (Section 3.2.3).

3.2.1 Plasma Physical Regime

The plasmas considered in this paper are collisionless, ultrarelativistically hot, and magnetized. This section discusses the parameters that characterize such a plasma. Throughout this work, we will use $\langle \cdot \rangle$ to denote the time-dependent volume- or particle ensemble-average of a quantity and $\bar{\cdot}$ to denote the time-average of a quantity in the time interval $10 < tc/L < 20$ unless otherwise specified; here L/c is the light-crossing time of the simulation domain with length L .

The parameters that characterize an ultra-relativistic, magnetized pair plasma with $\langle \gamma \rangle \gg 1$ include: an average total (electron plus positron) particle density n_0 , average particle energy $\langle \gamma \rangle m_e c^2$, and characteristic magnetic field strength $B_{\text{rms}} = \sqrt{\langle B_x^2 + B_y^2 + B_z^2 \rangle}$. Here, m_e is the mass of the electron (or positron) and $\langle \gamma \rangle$ is the average particle Lorentz factor ($\gamma = 1/\sqrt{1 - v^2/c^2} = \sqrt{1 + p^2/m_e^2 c^2}$ for a particle with velocity v , mass m_e , and momentum p). The fundamental physical length scales in this plasma are: the characteristic Larmor radius $\rho_e = \langle \gamma \rangle m_e c^2 / e B_{\text{rms}}$, the plasma skin depth $d_e = (\langle \gamma \rangle m_e c^2 / (4\pi n_0 e^2))^{1/2}$, and the size of the system L . For a plasma with a Maxwell-Jüttner particle distribution $f(\gamma) = \gamma^2 \sqrt{1 - 1/\gamma^2} [\theta K_2(1/\theta)]^{-1} \exp(-\gamma/\theta)$, the plasma has a well-defined temperature $T_e = \langle \gamma \rangle m_e c^2 / 3$ (assumed equal for electrons and positrons) and the Debye length simplifies to $\lambda_D = d_e / \sqrt{3}$. Here $\theta \equiv T_e / m_e c^2$ and $K_2(x)$ is the modified Bessel function of the second kind. The three length scales ρ_e , d_e , and L form two dimensionless quantities: the magnetization $\sigma = B_{\text{rms}}^2 / 4\pi h = 3(d_e / \rho_e)^2 / 4$ and the ratio of the largest characteristic scale of spatial variation $L/2\pi$ (which will be the turbulence driving scale in our study as described below) to the Larmor radius ρ_e . Here $h = n_0 \langle \gamma \rangle m_e c^2 + \langle P \rangle \approx 4n_0 \langle \gamma \rangle m_e c^2 / 3$ is the characteristic relativistic enthalpy density and $\langle P \rangle \approx n_0 \langle \gamma \rangle m_e c^2 / 3$ is the (assumed isotropic) average plasma pressure. The magnetization is related to the plasma beta parameter $\beta = 8\pi \langle P \rangle / B_{\text{rms}}^2$ as $\beta = 1/(2\sigma)$ and determines the relativistic Alfvén speed $v_A = c\sqrt{\sigma/(\sigma + 1)}$. Since we consider primarily Alfvénic turbulence, the magnetization governs how relativistic the large-scale turbulent motion is.

3.2.2 Numerical Simulations

To explore the properties of imbalanced turbulence in a collisionless relativistic pair plasma from first principles, we use the electromagnetic PIC code ZELTRON [109]. ZELTRON samples the particle phase space with macro-particles and evolves them according to the Lorentz force law, providing an approximate solution to the relativistic Vlasov equation. The electric and magnetic fields evolve according to Maxwell's equations, with the addition of an externally-driven volumetric current to Ampère's law to generate turbulence, as discussed below.

The physical parameters of the simulations we present are identical to those described in Ref. [128] and Ref. [114] except for the modifications outlined below to introduce imbalance. Each simulation is initialized with an electron-positron plasma at rest with a Maxwell-Jüttner distribution function, a uniform background magnetic field $\mathbf{B}_0 = B_0 \hat{\mathbf{z}}$ and no initial electromagnetic fluctuations. For each of the simulations, we set the initial magnetization to $\sigma_0 = 0.5$, yielding a relativistic Alfvén velocity $v_{A0} = 0.58c$ and plasma beta $\beta_0 = 1.0$. The initial temperature of the plasma is fixed at $T_e = 100m_e c^2$ across all simulations, corresponding to an initial average particle Lorentz factor of $\langle \gamma \rangle \approx 300$.

To obtain the largest possible inertial range, the simulation suite's chosen numerical parameters maximize the separation between the large driving scale $L/2\pi$ and the small initial kinetic scales ρ_{e0} and d_{e0} , while still resolving the latter. We resolve the initial plasma length scales with fixed $\Delta x = \rho_{e0}/1.5 = d_{e0}/1.22$, where $\Delta x = \Delta y = \Delta z$ is the grid cell length in each direction. The simulation domain is cubic with periodic boundary conditions and length $L \equiv N\Delta x$, where N is the number of cells in a spatial dimension (throughout, $N_x = N_y = N_z \equiv N$). The timestep is a fraction of the cell light-crossing time, i.e. $\Delta t = 3^{-1/2}\Delta x/c$. The simulations are initialized with 32 particles per cell per species. To scan the ratio $L/2\pi\rho_{e0}$, we vary the number of cells in each spatial dimension, with $N = 256, 384, 512$, and 768 corresponding to $L/2\pi\rho_{e0} \in \{27.1, 40.7, 54.3, 81.5\}$. When examining the dependence of results on simulation size, we also include three simulations of balanced turbulence with $L/2\pi\rho_{e0} \in \{81.5, 108.7, 164\}$ ($N \in \{768, 1024, 1536\}$) used in Ref. [114] that are otherwise identical to the simulations presented in this work.

The initial equilibrium is disrupted by an externally-driven current. We employ an oscillating Langevin antenna [129, OLA] to drive turbulence volumetrically and continually throughout each simulation's duration. The OLA is implemented by adding the external current to the evolution equation for the electric field (Ampère's law). This current generates counter-propagating Alfvén waves. The amplitudes of these counter-propagating waves are modified to induce imbalanced turbulence, as described in the following paragraphs. Because of the random nature of the OLA driving, a single simulation may not be representative of the entire ensemble of possible random seeds. To avoid basing all our conclusions on a single data point for each balance parameter, we also present a statistical study of random seeds. For each balance parameter, eight values of the random seed are simulated for the domain size $L/2\pi\rho_{e0} = 40.7$ ($N = 384$). The results of the statistical study are compared against the largest simulation domains $L/2\pi\rho_{e0} = 81.5$ ($N = 768$). Statistical variation could potentially be reduced in a single simulation by introducing more than eight driving modes.

We drive imbalanced turbulence via eight independently-evolved, externally-driven sinusoidal current modes. These current modes create magnetic field perturbations propagating in opposite directions along the background magnetic field, i.e. Alfvén waves. The driven current modes have the form:

$$J_x^{\text{ext}}(\mathbf{x}, t) = \frac{2\pi c}{L^2} \text{Re} \left[\sum_{j=1}^2 \left(a_j(t) e^{i\mathbf{k}_j \cdot \mathbf{x}} + b_j(t) e^{-i\mathbf{k}_j \cdot \mathbf{x}} \right) \right] \quad (3.1)$$

$$J_y^{\text{ext}}(\mathbf{x}, t) = \frac{2\pi c}{L^2} \text{Re} \left[\sum_{j=3}^4 \left(a_j(t) e^{i\mathbf{k}_j \cdot \mathbf{x}} + b_j(t) e^{-i\mathbf{k}_j \cdot \mathbf{x}} \right) \right] \quad (3.2)$$

$$J_z^{\text{ext}}(\mathbf{x}, t) = \frac{2\pi c}{L^2} \text{Re} \left[\sum_{j=1}^4 \left(-a_j(t) e^{i\mathbf{k}_j \cdot \mathbf{x}} + b_j(t) e^{-i\mathbf{k}_j \cdot \mathbf{x}} \right) \right]. \quad (3.3)$$

The sign of k_z dictates the direction of the current mode's propagation. Four of the modes have no y -component of their wavevector and four have no x -component; four propagate in the $+z$ -direction and four propagate in the $-z$ -direction. These wavevectors are

$$\mathbf{k}_1 = k_0(-1, 0, 1) \quad \mathbf{k}_2 = k_0(1, 0, 1) \quad (3.4)$$

$$\mathbf{k}_3 = k_0(0, -1, 1) \quad \mathbf{k}_4 = k_0(0, 1, 1). \quad (3.5)$$

Here, $k_0 = 2\pi/L$, so that the driving scale is the largest scale $L/2\pi$. We ensure $\nabla \cdot \mathbf{J}_{\text{ext}} = 0$ to avoid local injection of net charge. Currents driven in $J_{\text{ext},x}$ and $J_{\text{ext},y}$ create Alfvén waves with magnetic field

perturbations in the y - and x -directions, respectively. The amplitudes of these currents can be adjusted to create counter-propagating Alfvén waves of unequal amplitudes, thus enabling our study of imbalanced turbulence.

The external current's time-dependence is dictated by the coefficients $a_j(t)$ and $b_j(t)$. The coefficient at the $(n+1)$ th timestep is found from the previous n th timestep as

$$a_j^{(n+1)} = a_j^{(n)} e^{-i\omega\Delta t} + \alpha_j u_j^{(n)} \Delta t \quad (3.6)$$

$$b_j^{(n+1)} = b_j^{(n)} e^{-i\omega\Delta t} + \beta_j v_j^{(n)} \Delta t. \quad (3.7)$$

The coefficients $a_j(t)$ and $b_j(t)$ thus oscillate at frequency ω with random kicks at each timestep [cf. Langevin equation, hence the name ‘‘oscillating Langevin antenna’’; 129]. The initial coefficients $a_j^{(0)}$ and $b_j^{(0)}$ are set to amplitudes \mathcal{A} and \mathcal{B} multiplied by random phases $\phi_j^{(a)}$ and $\phi_j^{(b)}$: $a_j^{(0)} = \mathcal{A} e^{i\phi_j^{(a)}}$ and $b_j^{(0)} = \mathcal{B} e^{i\phi_j^{(b)}}$. We set $\mathcal{A} = B_0 L / 8\pi$, which for balanced turbulence achieves $\delta B_{\text{rms}} = \sqrt{B_{\text{rms}}^2 - B_0^2} \sim B_0$. The random kicks $u_j^{(n)}$ and $v_j^{(n)}$ in Equations 3.6 and 3.7 are complex random numbers with real and imaginary components drawn from a uniform distribution between -0.5 and 0.5 . The constant parameters α_j and β_j are set such that when ensemble-averaged, $\langle |a_j^{(n)}|^2 \rangle = \mathcal{A}^2$ and $\langle |b_j^{(n)}|^2 \rangle = \mathcal{B}^2$. The complex driving frequency ω has real component ω_0 and an imaginary component $-\Gamma_0$ which we set to be non-integer multiples of the Alfvén frequency $\omega_A \equiv 2\pi v_A / L$ to avoid initial resonances: $\omega_0 = (0.6/\sqrt{3})\omega_A \approx 0.35\omega_A$ and $\Gamma_0 = (0.5/\sqrt{3})\omega_A \approx 0.29\omega_A$. In frequency space, the driving is a Lorentzian centered at ω_0 with a full-width half-max of Γ_0 ; see Ref. [129] for details. The amplitudes \mathcal{A} and \mathcal{B} (and therefore the α_j and β_j values) are the same for all a_j and b_j , but the random parameters u_j , v_j , and the initial phases $\phi_j^{(a,b)}$ are different for each k_j .

We introduce imbalance by adjusting the amplitudes of the currents propagating in the $-z$ -direction relative to those propagating in the $+z$ -direction. The coefficients a_j control the currents propagating in the $+z$ -direction, whereas b_j control the waves propagating in the $-z$ -direction. These currents' amplitudes are dictated by their respective \mathcal{A} and \mathcal{B} amplitudes. To achieve imbalanced turbulence, we hold \mathcal{A} fixed and vary \mathcal{B} . We quantify how balanced the turbulence is via the *balance parameter*,

$$\xi \equiv \mathcal{B}/\mathcal{A}, \quad (3.8)$$

where \mathcal{B} is the amplitude for the $-z$ -modes and \mathcal{A} is the fixed amplitude for the $+z$ -modes. A value of $\xi = 1$ corresponds to the canonical balanced case, whereas $\xi = 0$ corresponds to current modes propagating only in the $+z$ -direction. Because we drive currents rather than Alfvén modes, we do not directly control the exact amplitude of counter-propagating Alfvén waves. If $\xi = 0$ corresponded exactly to the case of Alfvén waves propagating in a single direction, we would not expect turbulence to develop in a non-relativistic, ideal MHD plasma. Indeed, turning off the random kicks by setting $\alpha_j = \beta_j = 0$ does not result in turbulence for simulations with $\xi = 0$ (not shown). However, with our set-up of nonzero α_j and β_j , the $\xi = 0$ case does become turbulent because the OLA forcing excites counter-propagating Alfvén waves. In principle, more imbalanced turbulence should be achievable by, e.g., a decaying turbulence problem or by changing the driving mechanism. For the present work, we simply term the $\xi = 0$ case the “most imbalanced” case.

We will use the balance parameter ξ throughout this paper to refer to the degree of imbalance; however, since ξ measures the imbalance of the driving mechanism rather than the turbulence, we now briefly discuss the relationship of ξ to other methods of measuring imbalance. Cross-helicity, an invariant in ideal MHD, measures the difference in the energy densities associated with waves propagating anti-parallel (with energy density \mathcal{E}_+ and amplitude δB_+) and parallel to the magnetic field (with energy density \mathcal{E}_- and amplitude δB_-). In ideal, non-relativistic, incompressible MHD, $\mathcal{E}_\pm = \langle \rho |z_\pm|^2 \rangle / 4$, where $z_\pm = \delta \mathbf{v} \pm \mathbf{b}$ are the Elsasser fields [130]. Here $\delta \mathbf{v}$ is the fluctuating plasma velocity, $\mathbf{b} \equiv \delta \mathbf{B} v_A / B_0$ is the fluctuating magnetic field in velocity units, and ρ is the plasma mass density (not to be confused with the Larmor radius ρ_e). The total energy density is then given by $\mathcal{E} \equiv \mathcal{E}_+ + \mathcal{E}_- = (1/2) \langle \rho (|\delta \mathbf{v}|^2 + |\mathbf{b}|^2) \rangle$ and the cross-helicity $H_c \equiv (\mathcal{E}_+ - \mathcal{E}_-) / \langle \rho \rangle$ can be re-expressed as $H_c = \langle \delta \mathbf{v} \cdot \mathbf{b} \rangle$. Cross-helicity is related to the volume-averaged z -component of the Poynting flux $\mathcal{S}(\mathbf{x}, t) = (c/4\pi) [\mathbf{E} \times \mathbf{B}]$ under the assumptions of incompressible, nonrelativistic, ideal reduced MHD that $\delta B \ll B_0$ and that the fluctuations $\delta \mathbf{B}$ and $\delta \mathbf{v}$ are perpendicular to the background field $B_0 \hat{z}$:

$$\langle S_z \rangle(t) = -\frac{B_0}{4\pi} \langle \delta \mathbf{v} \cdot \delta \mathbf{B} \rangle = -\frac{1}{4\pi} \frac{B_0^2}{v_A} H_c(t). \quad (3.9)$$

For a single Alfvén wave, $\delta \mathbf{v} / v_A = \pm \delta \mathbf{B} / B_0$ and thus the magnitude of the Poynting flux for a single Alfvén

Table 3.1: Measured values of the Elsasser fields' energy ratio $r_E = \overline{\langle z_-^2 \rangle} / \overline{\langle z_+^2 \rangle}$ for a sampling of balance parameter ξ values. The first number gives the ratio for $N = 768$ and the second gives one standard deviation of the $N = 384$ statistical seed studies. Time averages are taken over $5.0 < tc/L < 20.0$.

ξ	0.0	0.5	1.0
r_E	0.72 ± 0.06	0.92 ± 0.14	1.01 ± 0.14

wave $|\langle S_z \rangle_{1\text{-wave}}|$ is

$$|\langle S_z \rangle_{1\text{-wave}}| = \frac{1}{4\pi} \langle \delta B^2 \rangle v_A. \quad (3.10)$$

We can estimate the values of the driven Alfvén wave energies in our simulations of imbalanced turbulence as $E_+ \sim \langle \delta B_+^2 \rangle \sim |a_j|^2$ (where δB_+ is the amplitude of the magnetic perturbation travelling in the $+z$ direction) and $E_- \sim |b_j|^2$, leading to

$$\langle S_z \rangle \propto H_c \propto 1 - \xi^2. \quad (3.11)$$

Equation 3.11 will be tested in Section 3.3.3.2. Normalizing to the total energy, the “normalized cross-helicity” $\tilde{H}_c = (\mathcal{E}_+ - \mathcal{E}_-) / (\mathcal{E}_+ + \mathcal{E}_-)$ [29, 45, 124] can be estimated as $\tilde{H}_c \sim (1 - \xi^2) / (1 + \xi^2)$. Finally, we calculate the ratio $\overline{\langle z_-^2 \rangle} / \overline{\langle z_+^2 \rangle}$ of Elsasser field energies for several ξ (Table 3.1). These fields are calculated with $v_A(t) = c\sqrt{\sigma(t)/(\sigma(t) + 1)}$ using the instantaneous magnetization. For our most imbalanced turbulence ($\xi = 0.0$), the ratio of energies is about 0.72, whereas for perfectly imbalanced turbulence the ratio would be zero. The discrepancy between the driving’s imbalance and the turbulence’s imbalance is due to the excitation of counter-propagating Alfvén waves (discussed in the previous paragraph), and possibly relativistic, kinetic, and moving frame effects, which the Elsasser fields we use do not take into account.

The main goal of our study is to determine the impact of imbalance on the properties of collisionless turbulence. We do so by varying the balance parameter ξ between 0 (“most imbalanced”) and 1 (“balanced”) at every value of $L/2\pi\rho_e0$.

3.2.3 Energy Diagnostics

In this section, we discuss diagnostics that will be used in Section 3.3 to partition the energy of the system into four main types.

The total energy density in the system can be decomposed into the energy density \mathcal{E}_{EM} in the electric and magnetic fields and the total (kinetic plus rest mass) energy density \mathcal{E}_{pl} in the plasma particles. Fluid quantities provide intuition into the plasma's behavior by partitioning \mathcal{E}_{pl} into internal, net flow, and turbulent flows:

$$\mathcal{E}_{\text{pl}}(\mathbf{x}, t) = \mathcal{E}_{\text{int}}(\mathbf{x}, t) + \mathcal{E}_{\text{net}}(t) + \mathcal{E}_{\text{turb}}(\mathbf{x}, t). \quad (3.12)$$

The plasma's total kinetic, internal, and turbulent kinetic energy densities are calculated for each simulation cell from both the electron and positron macro-particles' positions and momenta, though they will often be discussed in terms of their volume averages $\langle \mathcal{E}_{\text{pl}} \rangle$, $\langle \mathcal{E}_{\text{int}} \rangle$ and $\langle \mathcal{E}_{\text{turb}} \rangle$, respectively. The net flow energy density \mathcal{E}_{net} is a key quantity for characterizing how efficiently imbalanced turbulence can drive a directed plasma flow. It is a global quantity calculated from the total momentum in the system. Explicitly, these quantities are defined as:

$$\mathcal{E}_{\text{int}}(\mathbf{x}, t) \equiv \sqrt{\mathcal{E}_{\text{pl}}^2(\mathbf{x}, t) - \mathcal{P}_{\text{pl}}^2(\mathbf{x}, t)c^2} \quad (3.13)$$

$$\mathcal{E}_{\text{net}}(t) \equiv \langle \mathcal{E}_{\text{pl}} \rangle(t) - \sqrt{\langle \mathcal{E}_{\text{pl}} \rangle^2(t) - \langle \mathcal{P}_{\text{pl}} \rangle^2(t)c^2} \quad (3.14)$$

$$\mathcal{E}_{\text{turb}}(\mathbf{x}, t) \equiv \mathcal{E}_{\text{flow}}(\mathbf{x}, t) - \mathcal{E}_{\text{net}}(t) \quad (3.15)$$

where $\mathcal{E}_{\text{flow}}(\mathbf{x}, t) = \mathcal{E}_{\text{pl}}(\mathbf{x}, t) - \mathcal{E}_{\text{int}}(\mathbf{x}, t)$ [128, Eq. 8] and $\mathcal{P}_{\text{pl}}(\mathbf{x}, t)$ is the local momentum density of the electron-positron plasma. This framework is similar to that in Ref. [128], with the renaming of the ‘‘bulk’’ energy density to the ‘‘flow’’ kinetic energy density and further breakdown of the flow energy density into the energy density associated with the net flow of the plasma \mathcal{E}_{net} through the simulation domain and the turbulent motions $\mathcal{E}_{\text{turb}}$. Equation 3.13 for the internal energy density is analogous to the relativistic energy E of a single particle $E^2 = (mc^2)^2 + p^2c^2$. In this analogy, the plasma internal energy acts like a particle rest mass, the plasma momentum acts as a particle momentum, and the total plasma kinetic energy acts like a relativistic particle mass. Equation 3.14 for the net flow energy density uses a similar analogy, specifically applied to the quantity volume-averages.

The change in the energy of the plasma and the electromagnetic fields comes from the energy injected into the system by the OLA driving. The energy injection rate $\langle \dot{\mathcal{E}}_{\text{inj}} \rangle = -\langle \mathbf{J}_{\text{ext}} \cdot \mathbf{E} \rangle$ is statistically constant in time. Integrating it over time gives the total injected energy density up to time t : $\mathcal{E}_{\text{inj}}(t)$. Because the injected

energy depends on the amplitude of the driven waves, its value at any given time varies with ξ within a factor of two or so (see Section 3.3.2). To account for this dependence of injected energy on balance parameter, we normalize the volume-averaged changes in the various energy components by the volume-averaged injected energy to obtain energy efficiencies:

$$1 = \frac{\Delta\langle\mathcal{E}_{\text{int}}\rangle}{\langle\mathcal{E}_{\text{inj}}\rangle} + \frac{\Delta\mathcal{E}_{\text{net}}}{\langle\mathcal{E}_{\text{inj}}\rangle} + \frac{\Delta\langle\mathcal{E}_{\text{turb}}\rangle}{\langle\mathcal{E}_{\text{inj}}\rangle} + \frac{\Delta\langle\mathcal{E}_{\text{EM}}\rangle}{\langle\mathcal{E}_{\text{inj}}\rangle}, \quad (3.16)$$

where the terms on the right are the internal efficiency, net flow efficiency, turbulent kinetic efficiency, and electromagnetic efficiency, respectively. We use $\Delta\mathcal{E}$ to indicate the change in a type of energy density since the start of the simulation, i.e. $\Delta\mathcal{E}(t) = \mathcal{E}(t) - \mathcal{E}(0)$.

3.3 Results

In this section, we investigate how imbalanced turbulence differs from balanced turbulence through a series of comparisons. After demonstrating the presence of a turbulent cascade for all values of balance parameter (Section 3.3.1), we examine how the injected energy transforms into the plasma's internal and turbulent energy (Section 3.3.2). We next turn to the novel aspect of imbalanced turbulence: the presence of a net flow (Section 3.3.3). By using the statistical study of eight random seeds at smaller simulation domains $L/2\pi\rho_{e0} = 40.7$ to enhance the trends in the largest simulation domains $L/2\pi\rho_{e0} = 81.5$, we constrain the dependence of each of these energy types on balance parameter. We then explore the decomposition of the plasma energy into thermal and nonthermal components and how particle acceleration depends on the balance parameter (Section 3.3.4). The influence of simulation domain size on the fluid quantities is explored by varying $L/2\pi\rho_{e0}$ in Appendix 3.A.

3.3.1 Formation of a Turbulent Cascade

The spectrum of the turbulent magnetic energy is a common diagnostic when examining turbulence. Much of the previous work on imbalanced turbulence in MHD plasmas has examined the power-law indices of the two Elsasser fields and how they may or may not deviate from Goldreich-Sridhar $k_{\perp}^{-5/3}$ scalings [14,

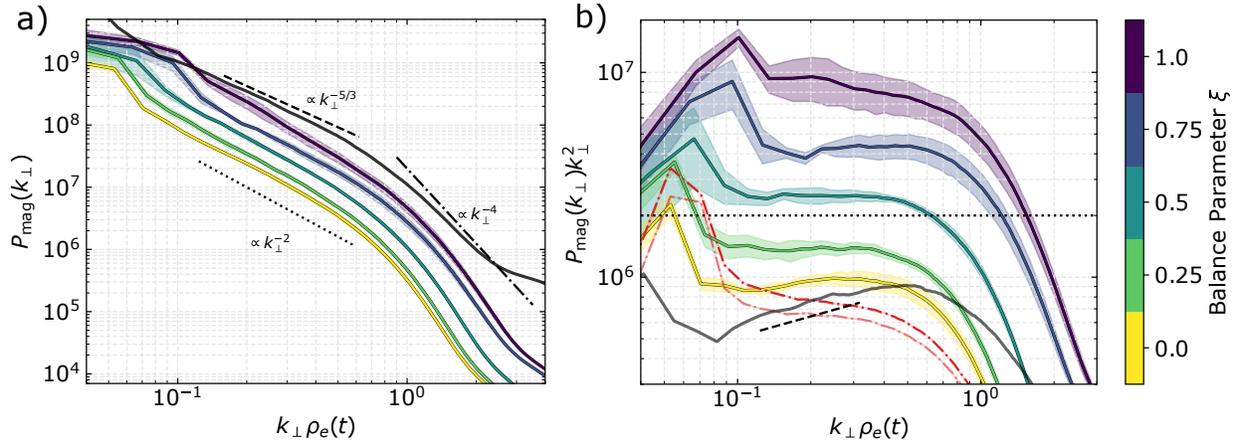


Figure 3.1: A turbulent cascade forms for all balance parameters. a) The magnetic energy spectra $P_{\text{mag}}(k_{\perp})$ for $L/2\pi\rho_{e0} = 81.5$ simulations of varying balance parameter averaged between times $8.8 < tc/L < 9.9$ (comprising five outputs) show an inertial range between $k_{\perp}\rho_e(t) \sim 0.1$ and 1.0 . A break in the spectrum at $k_{\perp}\rho_e(t) \sim 1.0$ indicates the onset of kinetic effects. b) When compensated by k_{\perp}^2 , the spectra for the balanced $\xi = 0.75$ and 1.0 cases are slightly steeper than $\propto k_{\perp}^{-2}$, whereas the imbalanced case $\xi = 0.0$ is slightly flatter. The Elsasser fields' spectra, shown in dash-dot red lines for $\xi = 0$, exhibit slightly different slopes, with the stronger field (z_+ , top line) being slightly steeper than the weaker field (z_- , bottom line). In both panels, shaded lines show one temporal standard deviation about the mean. Black dashed lines show the scaling $k_{\perp}^{-5/3}$; black dotted lines show k_{\perp}^{-2} . Gray lines show the $L/2\pi\rho_{e0} = 164$ balanced simulation's magnetic energy spectrum, taken at $t = 8.9 L/c$.

21, 24]. In this study, we simply calculate the overall turbulent magnetic energy spectrum via:

$$P_{\text{mag}}(k_{\perp}, t) = \int dk_z d\phi k_{\perp} \frac{1}{8\pi} \left(|\tilde{B}_x|^2 + |\tilde{B}_y|^2 + |\delta\tilde{B}_z|^2 \right), \quad (3.17)$$

where $\delta B_z = B_z - B_0$, k_z are the parallel wavenumbers, ϕ are the azimuthal angles, and $\tilde{\cdot}$ indicates the Fourier transform.

We find that the magnetic energy spectrum shows the formation of a turbulent cascade for all balance parameters (Fig. 3.1a). The spectra averaged over the time interval $8.8 < tc/L < 9.9$ (corresponding to $5.1 < tv_{A0}/L < 5.7$), i.e. after the turbulent cascade has fully developed but before the plasma's heating has diminished the inertial range, show similar shapes for all values of the balance parameter. The inertial range forms between $k_{\perp}\rho_e(t) \sim 0.08$ and 0.6 for the most imbalanced case; a slightly shifted inertial range beginning at $k_{\perp}\rho_e(t) \sim 0.15$ rather than $k_{\perp}\rho_e(t) \sim 0.08$ for the simulations with more balanced turbulence ($\xi = 1.0$ and 0.75) results from the faster heating at these balance parameters. The power-law index in the inertial range is roughly consistent with $-5/3$, the classic MHD prediction for strong turbulence [14]. Although the magnetic energy spectrum better matches the k_{\perp}^{-2} scaling characteristic of weak turbulence in the non-relativistic [131] and relativistic [132, 133] regimes, the turbulence in our simulations is strong. The steeper than $5/3$ spectrum is likely due to a small domain size, as found by Ref. [128]. Identical balanced simulations with twice the domain size are consistent with a power-law index of $-5/3$ (dark grey line in Fig. 3.1a). Accurately measuring the power-law indices of imbalanced turbulence via larger domain sizes is beyond the scope of the present study. Below the characteristic Larmor radius ($k_{\perp}\rho_e \gtrsim 1$), the spectrum steepens to another power law that covers a more limited range between $k_{\perp}\rho_e(t) \sim 1$ and 2 and is broadly consistent with the formation of a kinetic cascade with a power-law index of -4 for all values of imbalance [also found in 128], much steeper than in the inertial range. Again, however, providing more exact values to test against the predictions and measurements in Ref. [134] or Ref. [128] would require larger, better-resolved simulation domains. Numerical noise dominates at scales smaller than $k_{\perp}\rho_e(t) \sim 2$.

Although determining the precise dependence of the inertial range's slope on balance parameter would require a larger inertial range, there are hints that the slope depends on ξ . When the magnetic energy spectrum is compensated by the scaling k_{\perp}^{-2} , the simulation with balanced turbulence ($\xi = 1.0$) has

a downward-sloping spectrum, whereas the simulation of turbulence with $\xi = 0$ has a slightly positive slope (Fig. 3.1b). However, this steepening is most likely due to increased damping. The balanced simulations heat up faster (as discussed in Section 3.3.2), resulting in smaller values of L/ρ_e than the imbalanced simulations. Due to the small domain sizes, our simulations cannot distinguish differences in slope caused by imbalance or dissipation.

Because previous MHD predictions for imbalanced turbulence generally discuss the spectra of the Elsasser fields rather than the magnetic energy spectra, we also plot the Elsasser fields' spectra for the $\xi = 0$ case (dash-dot red lines in Fig. 3.1b). The larger-amplitude field appears to have a slightly steeper slope than the smaller-amplitude field, consistent with previous MHD simulations [19, 28]. However, constraining the variation in the power-law index is difficult to quantify with such a short inertial range. If the dependence of the Elsasser fields' energy spectra on imbalance persists in larger simulations, it could support MHD predictions that the spectra's power-law indices depend on imbalance [23, 131].

Further evidence for the formation of a turbulent cascade comes from comparing the evolution of the energy injected into the simulation and the internal energy of the plasma. Whereas the accumulated injected energy increases linearly from $t = 0$ onward (Fig. 3.2), the internal energy density does not begin to increase until about $2.5 L/c$, close to one Alfvén-crossing time (see Section 3.3.2.3, Fig. 3.3c). Presumably the energy injected at the driving scale cascades to smaller scales over the time period $t = 0 - 2.5 L/c$ until it reaches the characteristic Larmor radius and dissipates into internal energy—i.e. the turbulent cascade forms in the first couple of light-crossing times. There appears to be an increase in the cascade formation time for decreasing balance parameter (see Section 3.3.2.3).

3.3.2 Partition of the Injected Energy

3.3.2.1 Framework for the Energy Partition

Because we drive the plasma in each simulation without an energy sink, the overall energy of each case increases in time. By adding a statistically-constant amount of energy at each timestep, the overall amount of injected energy increases linearly in time (Fig. 3.2). The amplitude of the driven waves by definition

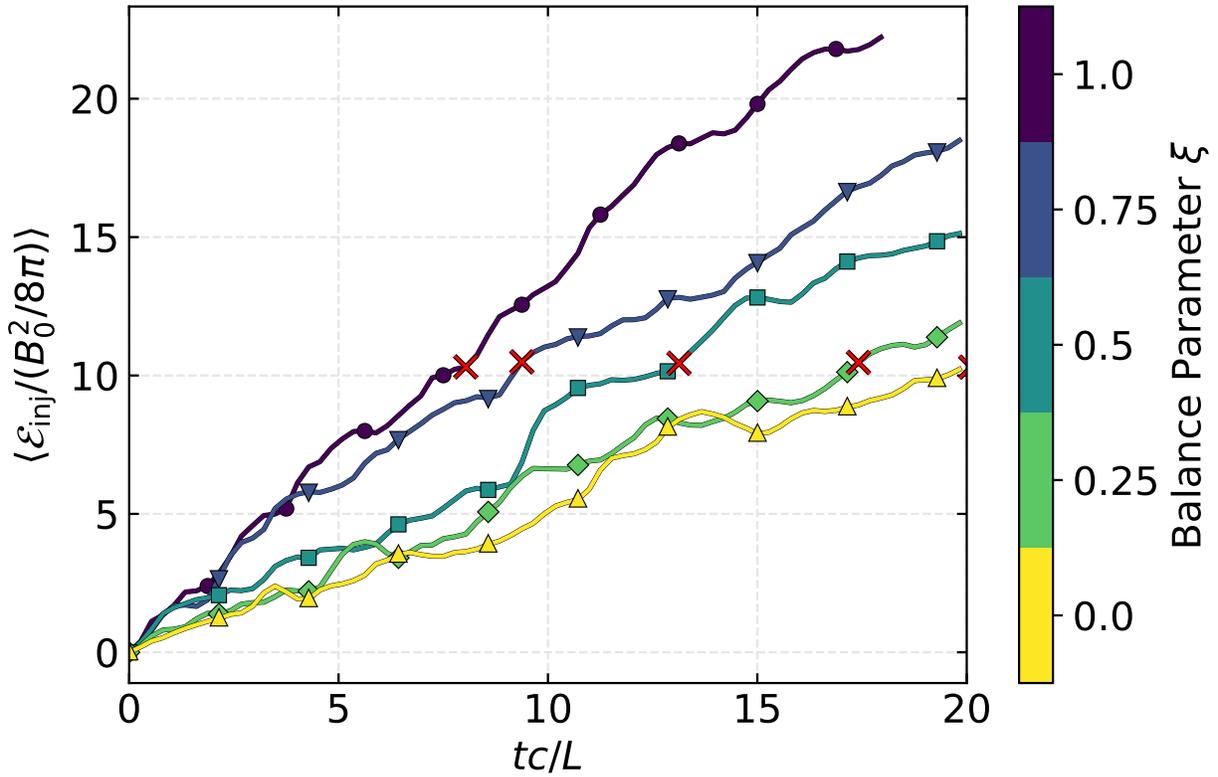


Figure 3.2: The amount of energy injected into a simulation depends on its balance parameter. The simulations of more balanced turbulence (purple and blue) have more injected energy than the simulations of less balanced turbulence (yellow and green). Red \times 's indicate the “equivalent” times where the same amount of energy has been injected for each simulation (see Table 3.3), which all have $L/2\pi\rho_{e0} = 81.5$.

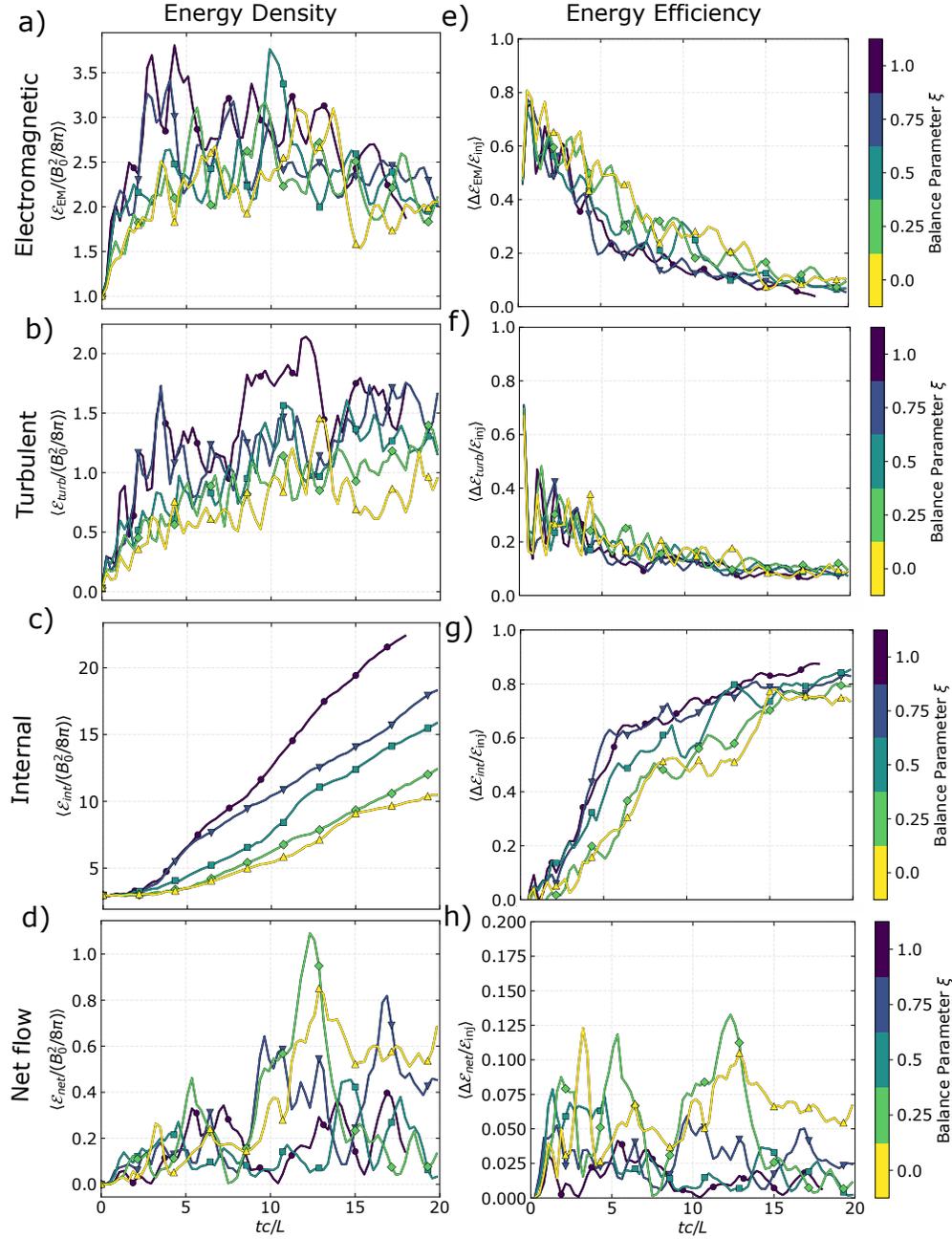


Figure 3.3: Energy partition into electromagnetic, turbulent kinetic, internal, and net flow energy depends on balance parameter. Left column: each type of energy density evolved over time, normalized to the constant value of the initial magnetic energy density $B_0^2/8\pi$. The turbulent electromagnetic (a) and kinetic (b) energy densities reach a constant value whereas the internal (c) and net flow (d) energy densities increase over time. Right column: the change in each type of energy density evolved over time, normalized to the total amount of injected energy density $\mathcal{E}_{\text{inj}}(t)$. Summing over the four panels on the right for each simulation adds to 1. Turbulent electromagnetic (e) and kinetic (f) energy efficiencies decay as $\propto t^{-1}$, whereas internal (g) and net flow (h) energy efficiencies saturate at a constant fraction of the injected energy. Note that the net flow energy (h) has a different vertical axis. Colors and markers indicate balance parameter. These simulations all have $L/2\pi\rho_{e0} = 81.5$.

depends on the balance parameter ξ (Equation 3.8), and so the amount of energy injected also depends on ξ . The increase in injected energy is twice as fast for the balanced case $\xi = 1.0$ as for the most imbalanced case $\xi = 0.0$ (Fig. 3.2). This doubling of injected energy occurs because twice as many modes are driven in the balanced as in the most imbalanced case.

The injected energy converts into various types of plasma energy, each of which will be discussed in the following subsections. Fig. 3.3 shows temporal evolution of each quantity's energy density (left) and energy efficiency (right). The large-scale injected energy cascades to smaller scales in the form of bulk kinetic and electromagnetic energy until it is dissipated into internal energy, thereby implying that the internal energy should increase linearly in time during the statistical steady state—and it does, as shown in Fig. 3.3c. As stationary conduits for the turbulent Alfvénic cascade, we expect both the turbulent kinetic and magnetic perturbations to fluctuate around steady-state values rather than continually increasing over time. Simulations support this idea of statistically steady-state values: the turbulent electromagnetic and kinetic energies saturate to their mean values around $7.5 L/c$ and fluctuate thereafter around these values (Fig. 3.3a, b). These values' dependence on balance parameter will be explored in Section 3.3.2.2, while the dependence of the internal energy density's slope on balance parameter is discussed in Section 3.3.2.3. The final component of plasma energy, the kinetic energy of net motion through the simulation domain, does not have an easily characterized evolution (Fig. 3.3d); its dependence on balance parameter will be discussed in Section 3.3.3.2.

Normalizing the change in each type of energy to the injected energy (the energy efficiency; Equation 3.16) allows direct comparison between turbulence with different balance parameters while accounting for the injected energy's dependence on ξ (Fig. 3.4). By the end of the simulations at $t = 20 L/c$, the percentage of injected energy that dissipates into internal energy depends on balance parameter, varying from $\approx 80\%$ for the most imbalanced case ($\xi = 0$) to $\approx 90\%$ for the balanced ($\xi = 1$) case. For all balance parameters, this percentage increases over time as the majority of injected energy converts into internal energy. In contrast, the fractions of injected energy that convert into turbulent electromagnetic and kinetic energy (i.e. the electromagnetic and kinetic energy efficiencies) decrease in time as $1/t$, plotted as dashed black lines in Fig. 3.4, with similar magnitudes for the balanced ($\xi = 1.0$) and most imbalanced ($\xi = 0.0$)

case. These fits are motivated by the discussion in the previous paragraph; when normalized to the injected energy $\propto t$, these two types of turbulent energy can be fit to the function $A + B/t$. The fraction of injected energy that converts into net flow energy differs by an order of magnitude between the balanced case (1%) and the imbalanced case (10%). In both cases, however, the net flow efficiency remains relatively constant in time, indicating that a constant fraction of injected energy converts into net flow energy, with a clear dependence on ξ (also seen in Fig. 3.3h).

With a broad framework for the temporal evolution of internal, turbulent kinetic, and magnetic energy in hand, the following sections will explore each type of energy's dependence on balance parameter using various averages and highlighting statistical variation with the random seed study.

3.3.2.2 Electromagnetic and Turbulent Kinetic Energy

After an initial transient period, the turbulent electromagnetic and kinetic energies become statistically constant in time (Fig. 3.3a, b). The turbulent electromagnetic energy increases until it contains approximately the same amount of energy as in the background field (Fig. 3.3a). As ξ increases from 0 to 1, $\langle \Delta \mathcal{E}_{\text{mag}} \rangle$ increases by about 50% for turbulence in the largest simulation domain sizes (Fig. 3.5a). In contrast, the electric energy density decreases from 10 - 25% of $B_0^2/8\pi$ for $\xi = 0.0$ to 10 - 17% for $\xi = 1.0$ (not shown). This decrease in electric energy density could be due to the decrease in the plasma velocity, which is approximately the Alfvén speed: $E \sim (\delta v/c) \times B \sim (v_A/c) B_0$. Faster heating in the balanced turbulence case leads to smaller v_A and hence smaller electric field. The simulations show that the time-averaged turbulent kinetic energy density $\langle \Delta \mathcal{E}_{\text{turb}} \rangle$ also depends on ξ , increasing from 75% of the background magnetic energy for the most imbalanced case ($\xi = 0.0$) to about 125% for balanced turbulence with $\xi = 1.0$ (Fig. 3.5b). The smaller values of $\langle \Delta \mathcal{E}_{\text{mag}} \rangle$ and $\langle \Delta \mathcal{E}_{\text{turb}} \rangle$ in the imbalanced case result from the injection of less energy in this case (Fig. 3.2). The fraction of injected energy that converts into turbulent and magnetic energy (i.e. the corresponding energy efficiencies) at any given time is almost independent of balance parameter (Fig. 3.3e, f), with slightly higher $\langle \Delta \mathcal{E}_{\text{EM}} \rangle / \langle \mathcal{E}_{\text{inj}} \rangle$ for $\xi = 0$ compared to $\xi = 1$.

To test whether the turbulence is Alfvénic, we use the ‘‘Alfvén ratio’’ $r_A \equiv \langle \Delta \mathcal{E}_{\text{turb}} \rangle / \langle \Delta \mathcal{E}_{\text{mag}} \rangle$. The Alfvén ratio is related to the residual energy E_r (defined as the difference between the turbulent kinetic

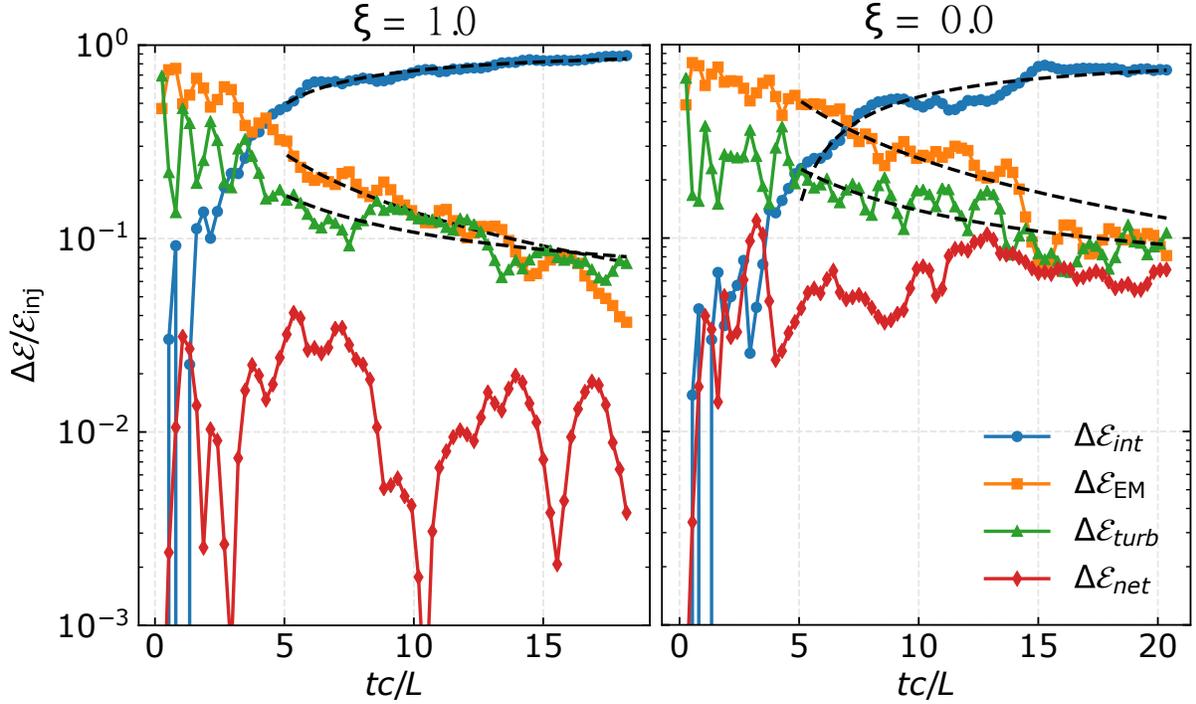


Figure 3.4: Time evolution of the energy partition for the balanced case ($\xi = 1.0$; left) and most imbalanced case ($\xi = 0.0$; right). Both show a decay in turbulent electromagnetic and kinetic energies and a saturation of internal and net flow energy densities; black dashed lines shows fits to $A + B/t$, with A and B constants. The imbalanced case has net flow energy density about an order of magnitude higher than the balanced case and correspondingly lower internal energy density, whereas the turbulent electromagnetic and kinetic energy densities are comparable for both balance parameters.

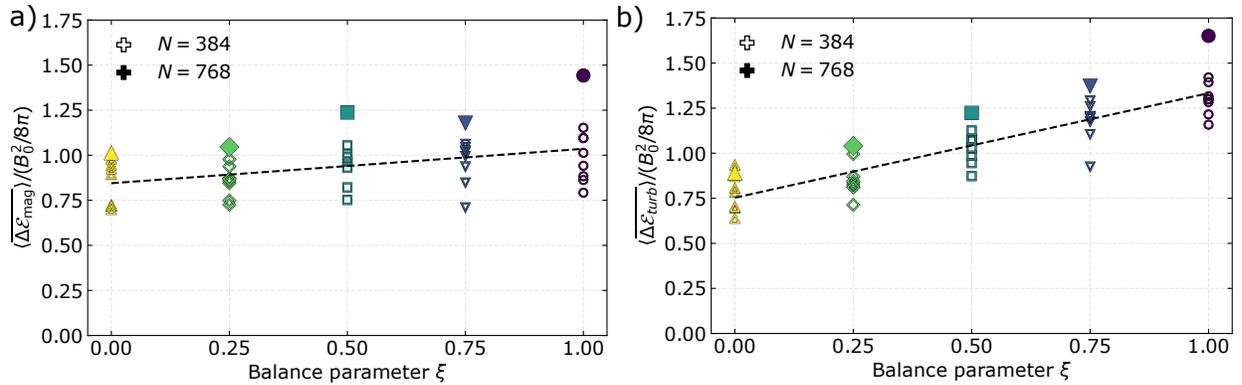


Figure 3.5: Trends of the turbulent magnetic and kinetic energy densities with balance parameter. Quantities are time-averaged from $10 < tc/L < 20$. The largest domain size $L/2\pi\rho_{e0} = 81.5$ (filled markers) shows a linear trend with balance parameter for the turbulent magnetic (a) and kinetic (b) energy densities, respectively. The statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). The dashed lines show linear fits. Colors and markers are the same as in Fig. 3.3.

and turbulent magnetic energies) via $E_r = (r_A - 1)/(r_A + 1)$. Ideal MHD predicts that the time- and volume-averaged kinetic and perturbed magnetic energies in an Alfvén wave (and thus perfectly imbalanced turbulence) should be in equipartition: $r_A = 1$. We might expect the same Alfvén ratio for turbulence (both balanced and imbalanced) comprising many Alfvén waves — though due to an increase in nonlinear interactions, physical plasmas such as the solar wind often have an excess of magnetic energy such that $r_A \approx 0.7$ [43, 45]; see Ref. [135], Ref. [136], and Ref. [137] for MHD models of this excess. Our simulations show that both balanced and imbalanced turbulence are in equipartition to within error bars (Table 3.2). The standard deviations of mean values for $\langle \overline{\Delta \mathcal{E}_{\text{turb}}} \rangle$ and $\langle \overline{\Delta \mathcal{E}_{\text{mag}}} \rangle$ in Table 3.2 were calculated for the $L/2\pi\rho_{e0} = 81.5$ and $L/2\pi\rho_{e0} = 40.7$ simulations (nine values for each balance parameter) and summed in quadrature. These error bars suggest that the large-scale ratio of turbulent kinetic to magnetic energies is independent of ξ . However, the residual energy may have a scale-dependent power-law spectrum with significant dependence on imbalance, which we do not address here. The solar wind shows a clear dependence of the residual energy spectrum’s slope on imbalance, with a value of -2 for balanced turbulence and closer to -1.8 for totally imbalanced turbulence [43]. The dependence of the residual energy spectrum’s slope on imbalance is a major outstanding puzzle that has not been successfully addressed by any phenomenological model of imbalanced turbulence thus far.

Our finding of approximate equipartition indicates that the turbulence is predominantly Alfvénic. In addition to Alfvén waves, slow and fast compressive modes also contribute to the turbulence. The fast modes, introduced by the OLA driving [92] or nonlinear relativistic wave conversion [138], and the slow modes, passively mixed by the turbulence [139], lead to total density fluctuations on the order of 20-30% of the background density in our simulations (not shown). Though the presence of fast and slow modes could affect the Alfvén ratio, characterizing their contribution is beyond the scope of this study.

3.3.2.3 Internal Energy

The increase in internal energy dominates the plasma energy budget at late times. Though the initial internal energy starts out at about three times the initial magnetic energy for all balance parameters, the plasmas with balanced turbulence heat up almost twice as quickly as the plasmas with imbalanced

Table 3.2: The turbulence in all simulations of balanced and imbalanced turbulence is approximately Alfvénic. The Alfvén ratio $r_A = \langle \Delta \mathcal{E}_{\text{turb}} \rangle / \langle \Delta \mathcal{E}_{\text{mag}} \rangle$ is approximately 1 for the largest simulations ($L/2\pi\rho_{e0} = 81.5$) for all values of the balance parameter ξ . Standard deviations are calculated from the statistical seed studies at each balance parameter.

ξ	$\langle \Delta \mathcal{E}_{\text{turb}} \rangle / \langle \Delta \mathcal{E}_{\text{mag}} \rangle$
1.0	1.1 ± 0.3
0.75	1.2 ± 0.2
0.5	1.0 ± 0.1
0.25	1.0 ± 0.1
0.0	0.9 ± 0.1

turbulence (Fig. 3.3c). At late times, about 80% of the injected energy is converted into internal energy in the imbalanced case ($\xi = 0$), slightly lower than the corresponding value of closer to 90% for the balanced case (Fig. 3.3g).

Because the plasma’s internal energy increases linearly after an initial transient, we characterize its heating rate (i.e. slope) through the dimensionless “injection efficiency” parameter η_{inj} . We define this order-unity coefficient as the ratio of the plasma heating rate $\langle \dot{\mathcal{E}}_{\text{int}} \rangle(t)$ to a “reference” heating rate $\langle \dot{\mathcal{E}}_{\text{ref}} \rangle$. We define $\langle \dot{\mathcal{E}}_{\text{ref}} \rangle$ by dividing the turbulent magnetic energy density, $\delta B_{\text{rms}}^2/8\pi$, by a characteristic non-linear cascade time at the outer scale $L/\delta v_{\text{rms}} \sim L/(v_{A0}\delta B_{\text{rms}}/B_0)$, assuming strong Alfvénic turbulence and $v_{A0}/c \ll 1$. This formulation gives us an operational definition of the injection efficiency in terms of quantities that can be directly measured in our simulations at any moment of time:

$$\eta_{\text{inj}} \equiv \frac{\langle \dot{\mathcal{E}}_{\text{int}} \rangle}{\langle \dot{\mathcal{E}}_{\text{ref}} \rangle} = \frac{8\pi B_0 L \langle \dot{\mathcal{E}}_{\text{int}} \rangle}{\delta B_{\text{rms}}^3 v_{A0}}. \quad (3.18)$$

The injection efficiency quantifies how efficiently turbulent magnetic energy cascades to small scales and dissipates. The heating rate of the plasma is extracted by fitting the slope of the internal energy in the interval $5 < tc/L < 20$ and converting to the injection efficiency via Equation 3.18, taking the value of δB_{rms} as the time-average over the same time period. The statistical mean value of the injection efficiency varies from about 1.0 for the most imbalanced case ($\xi = 0.0$) to about 1.5 for the balanced case ($\xi = 1.0$) as seen in Fig. 3.6a, with a statistical standard deviation on the order of 0.2.

As discussed in Section 3.3.2.2, the magnitude of the magnetic field fluctuations depends on balance parameter, which in principle may influence energy dissipation and the injection efficiency. To verify that

the trend in the injection efficiency η_{inj} with balance parameter ξ is not due to the variation of the amplitude of magnetic energy perturbations $\delta B_{\text{rms}}^2/B_0^2$ with ξ , we ran a simulation of imbalanced turbulence with $\xi = 0$ and a driving amplitude $\sqrt{2}$ times its canonical value. The magnetic field fluctuation level δB_{rms}^2 changes from $0.8B_0^2$ for the unadjusted amplitude case to $1.2B_0^2$ for the increased amplitude case, consistent with the unadjusted amplitude of the balanced case with the same random seed. The injection efficiency, however, increases to 1.1 with the increased amplitude, compared to 0.9 for the unadjusted $\xi = 0$ case and 1.4 for the unadjusted $\xi = 1$ case with the same random seed. Thus, two simulations with the same level of magnetic field perturbations but different balance parameters experience significantly different injection efficiencies, suggesting that the injection efficiency has an inherent dependence on balance parameter rather than δB_{rms} . This result suggests that the cascade time depends on balance parameter.

To explore the possibility that the cascade time τ_{casc} depends on balance parameter, we define

$$\tau_{\text{casc}} \equiv \frac{\langle \Delta \mathcal{E}_{\text{EM}} + \mathcal{E}_{\text{turb}} \rangle}{\langle \dot{\mathcal{E}}_{\text{int}} \rangle}. \quad (3.19)$$

The cascade time is normalized to the global Alfvén time $L/v_A(t)$ and then time-averaged over $10 < tc/L < 20$ (i.e. $5.8 < tv_{A0}/L < 11.6$). We expect the cascade time to be on the order of an Alfvén time for an Alfvénic cascade, and for balanced turbulence we indeed find that the cascade time varies statistically between 0.8 and 1.2 $L/v_A(t)$ (Fig. 3.6b). However, the cascade time increases to about 2.2 $L/v_A(t)$ on average for the imbalanced case $\xi = 0$. The lack of overlap between the cascade times of imbalanced and balanced turbulence suggest that the difference is statistically significant, rather than a fluke of random seeds. A longer cascade time for more imbalanced turbulence is consistent with Ref. [21]’s suggestion that the dominant waves are less strongly scattered in the imbalanced case.

3.3.3 Net Flow Energy and Momentum Transfer

As a component of the energy not present in balanced turbulence, we expect the kinetic energy in the net motion of the plasma through the simulation domain to depend on the balance parameter. In the balanced case, the statistically-symmetric (although not necessarily momentum-conserving) driving should on average lead to no net motion. In contrast, because imbalanced driving breaks the symmetry along the

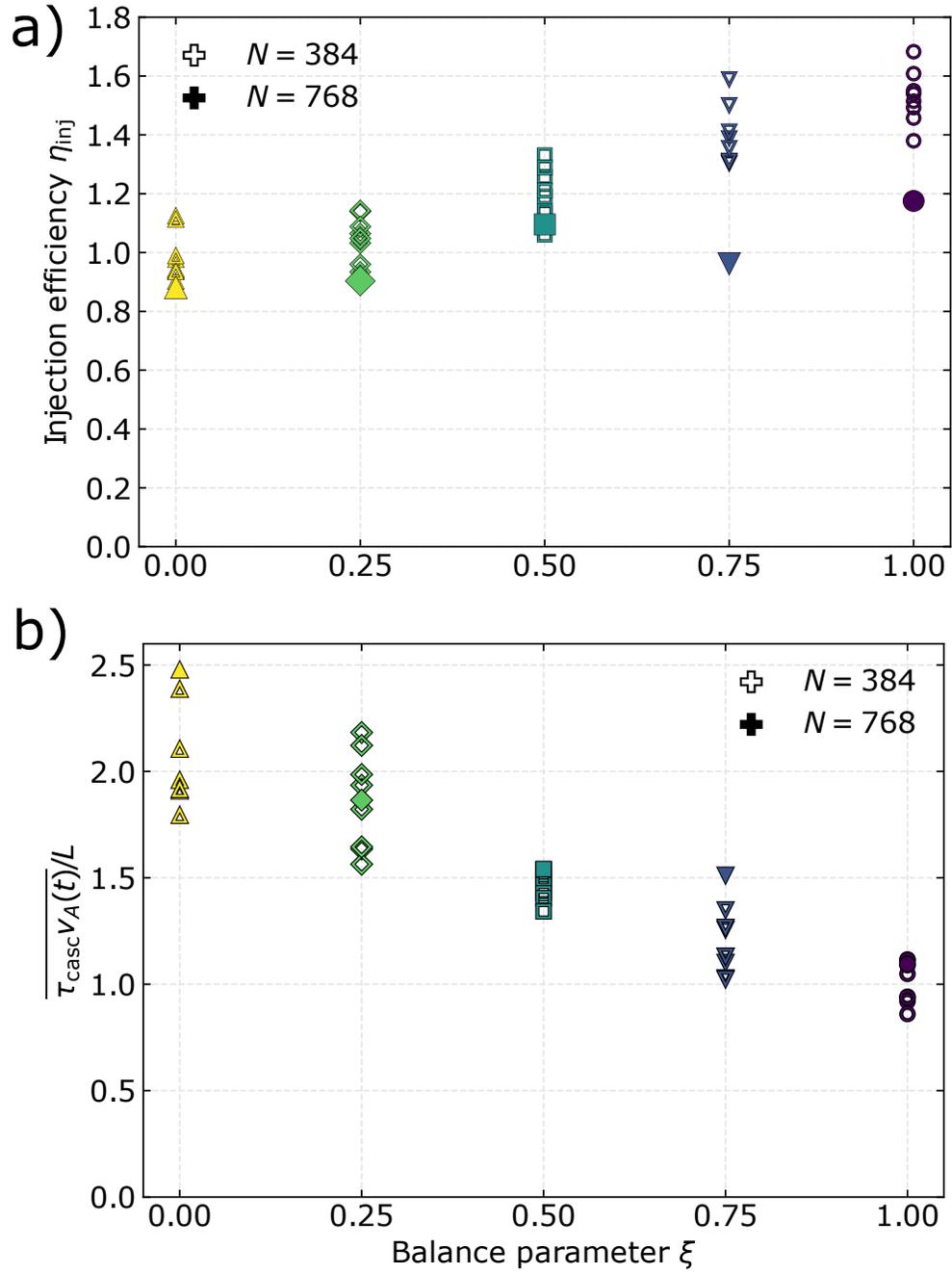


Figure 3.6: The injection efficiency η_{inj} (a; Equation 3.18) and cascade time τ_{casc} (b; Equation 3.19) depend on balance parameter. The largest domain size $L/2\pi\rho_{e0} = 81.5$ is shown with filled markers and the statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). Colors and markers are the same as in Fig. 3.3.

background magnetic field, we may expect a nonzero net flow energy for $\xi < 1.0$ if the asymmetric wave momentum converts into plasma momentum. In a gravitational potential, the net flow that results from efficient wave-plasma momentum coupling could form a wind or outflow. In this section, we first propose a simple model for the properties of such a net flow (Section 3.3.3.1) and then compare to numerical results (Section 3.3.3.2).

3.3.3.1 Analytic Framework for Momentum Transfer

A net flow could result from the efficient transfer of injected wave momentum into plasma momentum. In this section, we propose that the net flow velocity should be constant and provide a scaling for its magnitude. We can write the net flow energy density as

$$\mathcal{E}_{\text{net}}(t) = (\Gamma_{\text{net}}(t) - 1) \langle \rho \rangle(t) c^2 \approx \frac{\langle \mathcal{E}_{\text{int}} \rangle(t) v_{\text{net}}^2(t)}{2c^2}, \quad (3.20)$$

where the last expression holds for $v_{\text{net}}(t) \ll c$. We have defined the net Lorentz factor $\Gamma_{\text{net}}(t) = (1 - v_{\text{net}}^2/c^2)^{-1/2}$.

Analogously, the internal energy density relates to the net plasma momentum density as

$$\langle \mathcal{P}_{z,\text{tot}} \rangle(t) = \Gamma_{\text{net}}(t) \langle \rho \rangle(t) v_{\text{net}}(t) \approx \frac{\langle \mathcal{E}_{\text{int}} \rangle(t)}{c^2} v_{\text{net}}(t), \quad (3.21)$$

where again $v_{\text{net}} \ll c$ in the last expression. In simulations, both $\mathcal{E}_{\text{net}}(t)$ and $\langle \mathcal{P}_{z,\text{tot}} \rangle(t)$ increase linearly in time (Section 3.3.3.2). Because these two quantities depend on different powers of v_{net} , we deduce that the net velocity should be relatively constant in time. We test this prediction in the next section.

We can understand the net flow as a relativistic effect. As a limiting case, assume that the maximal asymmetric Poynting flux for a single Alfvén wave $|\langle S_z \rangle_{1\text{-wave}}|$ (Equation 3.10) is injected into a plasma. If a fraction ϵ of the momentum density $|\langle S_z \rangle_{1\text{-wave}}|/c^2$ in this electromagnetic wave converts into the plasma momentum density $\langle \mathcal{P}_{z,\text{tot}} \rangle$, we have

$$\langle \mathcal{P}_{z,\text{tot}} \rangle = \epsilon \frac{\langle \delta B^2 \rangle v_{A0}}{4\pi c^2} = \Gamma_{\text{net}} \langle \rho \rangle v_{\text{net}}, \quad (3.22)$$

where the last equality follows from Equation 3.21. By writing the relativistic mass density $\langle \rho \rangle = \langle \gamma \rangle m_e n_0$,

we find

$$\Gamma_{\text{net}} v_{\text{net}} = \frac{\langle \delta B^2 \rangle}{4\pi \langle \rho \rangle c^2} \epsilon v_{A0} = \frac{4}{3} \epsilon \delta \sigma v_{A0} \quad (3.23)$$

$$v_{\text{net}} = c \epsilon \sqrt{\frac{16 \delta \sigma^2 v_{A0}^2 / c^2}{9 + 16 \delta \sigma^2 v_{A0}^2 / c^2}} \rightarrow \frac{4}{3} \epsilon \delta \sigma v_{A0} \quad (3.24)$$

where we have defined $\delta \sigma = \langle \delta B^2 \rangle / (4\pi h)$ and the last expression again holds for $v_{\text{net}} \ll c$. From Equation 3.24, we see that the net flow's velocity approaches 0 as the magnetization goes to 0. Though the efficiency ϵ of converting electromagnetic momentum into plasma momentum can never be greater than 1, it could change with σ .

3.3.3.2 Numerical Results for Momentum Transfer

In this section, we test the assumptions behind the above calculations and demonstrate that our simulations do indeed find a net flow in line with the above framework.

Our simulations find that about 8% of the injected energy converts into net flow energy even in the most imbalanced case ($\xi = 0.0$), a fraction that becomes comparable to the turbulent electromagnetic or kinetic energy efficiencies after the latter two have decayed by a factor of two or more, around $15 L/c$ (Fig. 3.4). Since the net flow efficiency fluctuates strongly about a mean value in time, we compare different balance parameters by averaging over $10 < tc/L < 20$. The result, shown in Fig. 3.7, reveals that the net flow efficiency increases with decreasing ξ , as expected. The mean over an ensemble of identical simulations with $\xi = 0$ and varying random seeds is about 0.06, approximately three times as large as that for the balanced turbulence (about 0.02). The statistical spread in the balanced case is on the order of 0.01, and about 0.03 for the imbalanced case. Because energy is a strictly positive quantity and a limited number of modes were driven, even the simulations of balanced turbulence ($\xi = 1.0$) have a non-zero (albeit small) net flow energy due to short periods of net motion. Notably, there are three prominent outliers in Fig. 3.7 in the magnitude of the net flow efficiency in simulations with balance parameters $\xi = 0.5, 0.75$, and 1.0 . Each of these outlying simulations was initialized with the same random seed, suggesting that the particular random phases of the driven modes resulted in non-zero mean velocity later in the simulations' evolutions. The presence of these outliers demonstrates the need for a suite of random seeds to tease out statistically-robust trends.

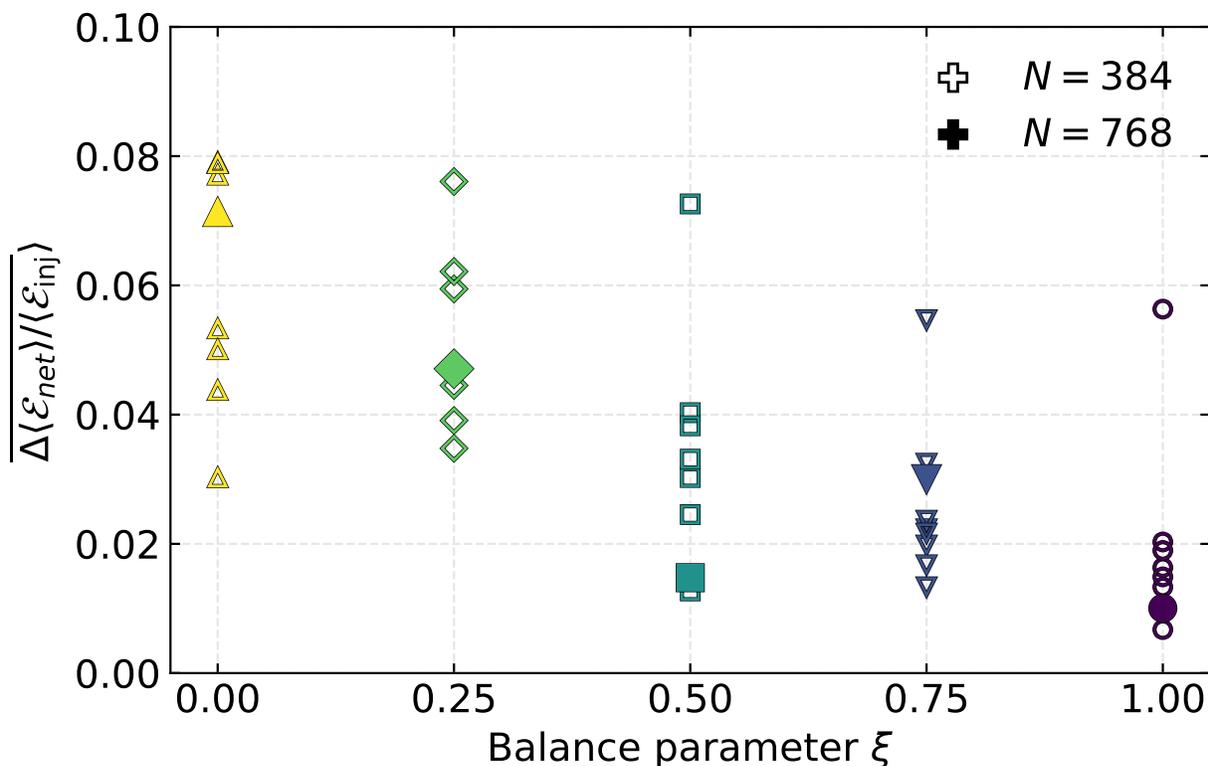


Figure 3.7: The net flow energy efficiency decreases with increasing balance parameter. The plotted values are volume-averaged and time-averaged from $10 < tc/L < 20$. The largest domain size $L/2\pi\rho_{e0} = 81.5$ is shown with filled markers and the statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). Note that the outliers for $\xi = 0.5, 0.75$, and 1.0 with net flow energy efficiencies a factor of 2 higher than the rest of the seed study were run with the same random seed. Colors and markers are the same as in Fig. 3.3.

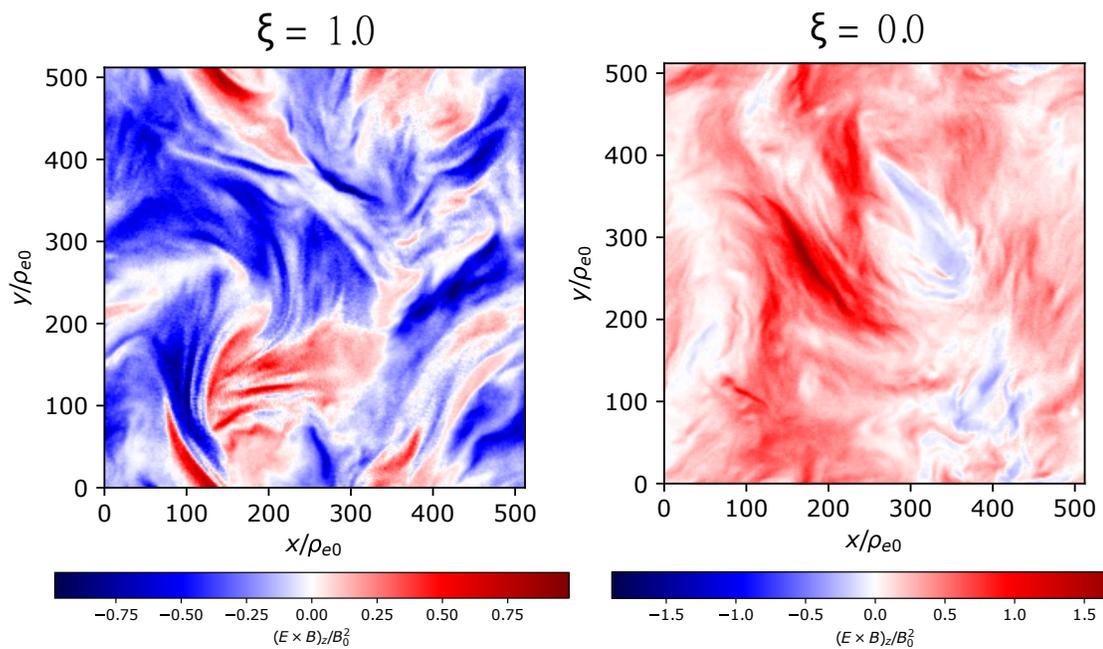


Figure 3.8: Even turbulence that is balanced as a whole has spatial and temporal pockets of locally imbalanced turbulence. Slices of the Poynting flux $(c/4\pi)\mathbf{E} \times \mathbf{B}$ in the z -direction, taken at the plane $z = 0$ at time $t = 16.1 L/c$ and normalized to $(c/4\pi)B_0^2$ for balanced turbulence ($\xi = 1.0$; left) and most imbalanced turbulence ($\xi = 0.0$; right) show variation in the sign of Poynting flux throughout the domain.

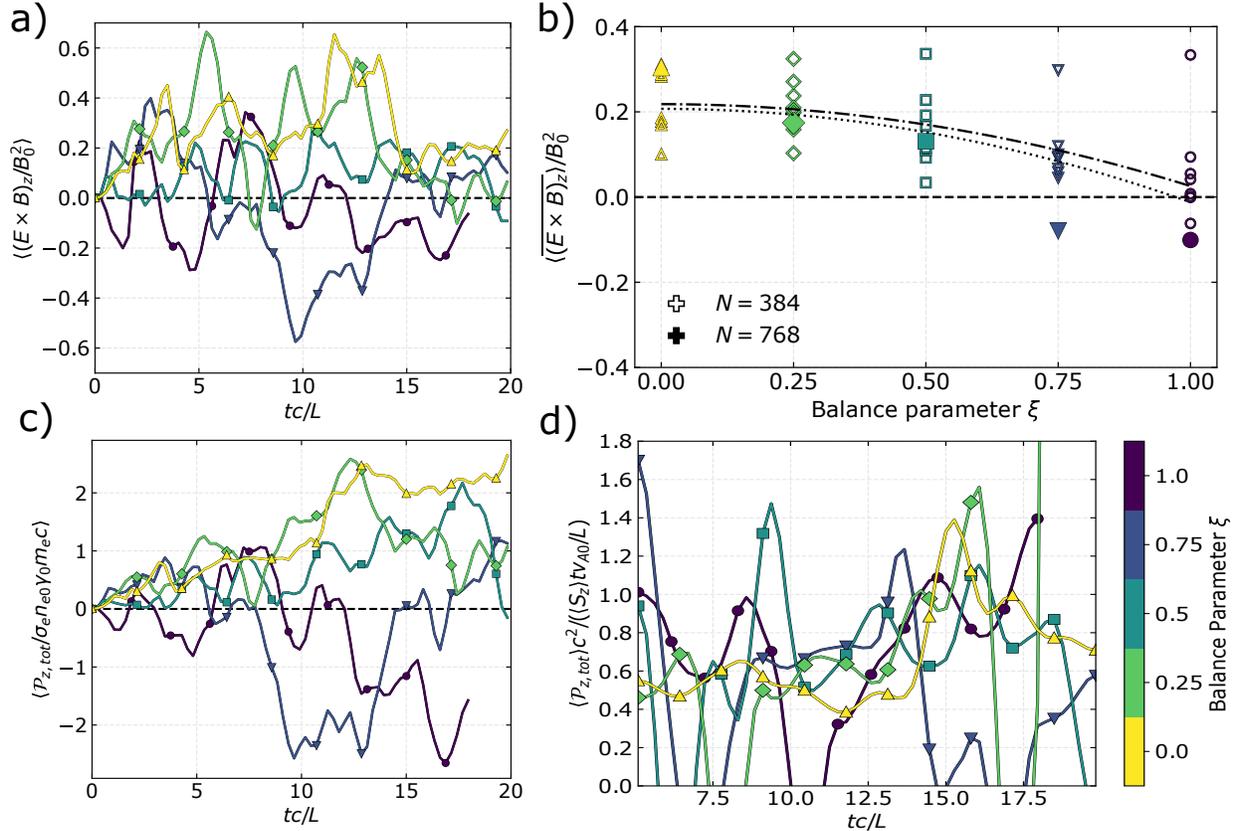


Figure 3.9: The average parallel Poynting flux is approximately constant in time, whereas the z -momentum of the plasma increases linearly in time. a) The time evolution of the volume-averaged Poynting flux $(c/4\pi)\langle \mathbf{E} \times \mathbf{B} \rangle$ in the z -direction, normalized to $(c/4\pi)B_0^2$, shows fluctuations around some mean value; time-averaging the curves over $10 < tc/L < 20$ shows a quadratic dependence on balance parameter (b). The time evolution (c) of the parallel plasma momentum shows an increase in time. The ratio $\langle \mathcal{P}_{z,\text{tot}} \rangle^2 / (\langle S_z \rangle tv_{A0}/L)$, shown in (d), is of order unity for all values of balance parameter. Black dash-dot lines show a quadratic fit; the dotted line is a quadratic fit without the outlier seed. Dashed black lines indicate zero. Colors and markers are the same as in Fig. 3.3.

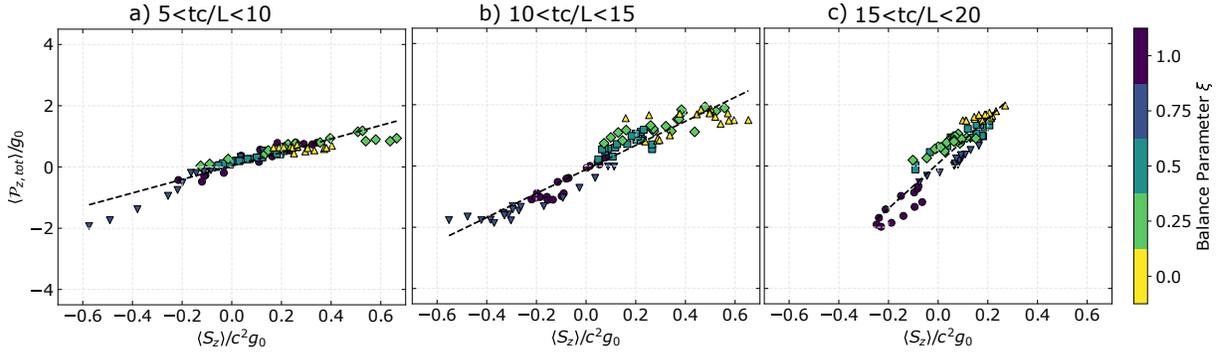


Figure 3.10: The parallel electromagnetic momentum generates and maintains the parallel momentum of the plasma. In these plots, each point corresponds to the instantaneous values of the volume-averaged plasma momentum along the background magnetic field and the volume-averaged parallel electromagnetic momentum at a given simulation time. Three periods of time are shown with linear fits of slopes 2.2, 3.9, and 7.4 for (a) $5 < tc/L < 10$, (b) $10 < tc/L < 15$, and (c) $15 < tc/L < 20$, respectively. Here, $g_0 = B_0^2/(4\pi c)$ is a typical value of momentum density. Colors and markers are the same as in Fig. 3.3.

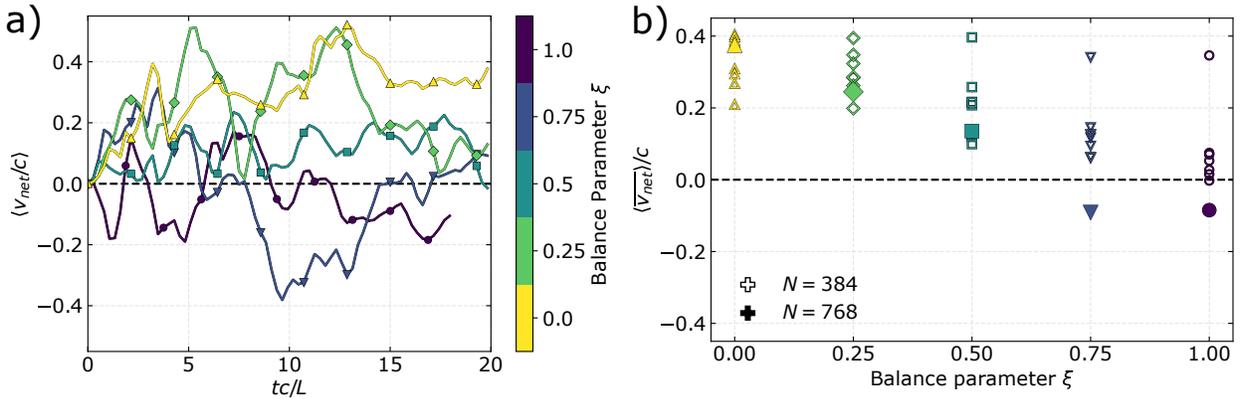


Figure 3.11: The net plasma velocity along the background magnetic field depends on balance parameter. a) The time evolution of the volume-averaged velocity v_{net} (Equation 3.21) shows values that fluctuate in time around some mean value that depends on ξ ; time-averaging the curves over $10 < tc/L < 20$ shows a dependence on balance parameter (b). The largest domain size $L/2\pi\rho_{e0} = 81.5$ is shown with filled markers and the statistical deviation is shown by the $L/2\pi\rho_{e0} = 40.7$ seed study (unfilled markers). Colors and markers are the same as in Fig. 3.3. Black dashed lines show zero.

Although the net flow energy density demonstrates the importance of the net flow in the overall energy budget, it does not contain information about the direction of the plasma's net motion. To address this issue, we now examine momentum rather than energy. First, we discuss the injected electromagnetic momentum and look at the Poynting flux $\mathbf{S} = (c/4\pi)\mathbf{E} \times \mathbf{B}$. For balance parameters $\xi < 1.0$, we expect the driven waves' Poynting flux to be nonzero along the background magnetic field ("parallel Poynting flux"; S_z). The spatial distribution of the parallel Poynting flux, plotted in Fig. 3.8, is highly nonuniform. Similar to MHD turbulence [29], our balanced simulation has local patches of strong Poynting flux and thus high imbalance, highlighting the fundamental connection between balanced and imbalanced turbulence at small scales. The total, volume-averaged parallel Poynting flux $\langle S_z \rangle$ is statistically constant in time and, for the balanced case ($\xi = 1.0$), oscillates around zero (Fig. 3.9a). After time-averaging over the period from $10 < tc/L < 20$, the parallel Poynting flux is clearly positive for imbalanced turbulence and consistent with zero for balanced turbulence (Fig. 3.9b). The value $\langle \overline{S_z} \rangle \approx 0.2 (cB_0^2/4\pi)$ at $\xi = 0.0$ is about 30% of the limiting value $(\delta B/B_0)^2 v_A/c (cB_0^2/4\pi) \approx v_A/c (cB_0^2/4\pi) \approx 0.58 (cB_0^2/4\pi)$ expected for a single Alfvén wave (Equation 3.10). The decrease in the Poynting flux with increasing balance parameter agrees well with the quadratic fit predicted by Equation 3.11, as shown by the dotted and dash-dot black lines in Fig. 3.9b, though we do not rule out a linear dependence. The same outliers discussed for the net flow efficiency are present in the parallel Poynting flux. The time- and volume-averaged Poynting flux in the x - and y -directions are much smaller than the parallel Poynting flux: within $0.1 (cB_0^2/4\pi)$ of zero for all balance parameters (not shown), indicating that the net electromagnetic momentum is primarily along the background magnetic field.

The injected Poynting flux imparts net momentum to the plasma. In agreement with the interpretation in Section 3.3.2.1, the volume-averaged parallel momentum $\langle \mathcal{P}_{z,\text{tot}} \rangle$ of the plasma increases approximately linearly over time (Fig. 3.9c). The ratio $\langle \mathcal{P}_{z,\text{tot}} c^2 \rangle / (\langle S_z \rangle tv_{A0}/L)$, shown in Fig. 3.9d, is approximately constant in time and fluctuates around values of order unity for any given balance parameter. In this figure, the parallel Poynting flux has been converted to an electromagnetic momentum density $\langle S_z \rangle / c^2$, and the momentum densities of both the plasma and the electromagnetic waves are normalized to $g_0 \equiv B_0^2 / (4\pi c)$, allowing for direct comparison between the two quantities. The ratio of the volume averages has been

smoothed with a Hanning window to remove arbitrarily large values due to a small Poynting flux. To further illustrate the relationship between $\langle S_z \rangle$ and $\langle \mathcal{P}_{z,\text{tot}} \rangle$, Fig. 3.10 shows their values at each time snapshot (given by individual dots) for each balance parameter (shown via marker and color). In general, the plasma momentum dominates over the electromagnetic momentum, with a time-dependent ratio given by the slope of the linear fit. The positive ratio shows that a positive volume-averaged parallel electromagnetic momentum corresponds to a positive parallel plasma momentum, as expected if Poynting flux converts to plasma momentum. The balanced turbulence's electromagnetic momentum and plasma momentum span both positive and negative values; for more imbalanced turbulence, the distribution shifts up and right, demonstrating that the asymmetric driving of electromagnetic momentum results in asymmetric net motion of the plasma in the z -direction.

Using Equation 3.21 to solve for v_{net} shows that v_{net} fluctuates around a mean value dependent on balance parameter (Fig. 3.11a). The net velocity of the plasma with balanced turbulence oscillates around zero, never reaching more than $0.2c$. The plasmas with the most imbalanced turbulence can reach net velocities up to $0.5c$. Averaging over $10 < tc/L < 20$ shows a clear dependence of the net velocity on balance parameter (Fig. 3.11b). As expected, plasmas with balanced turbulence experience a time-averaged net velocity near zero, though the finite simulation duration means that temporary movements parallel or anti-parallel to the background magnetic field are not completely averaged out. Equation 3.24 predicts a net velocity of $0.4c$ for the most imbalanced turbulence (plugging in four waves with $\delta B \sim B_0$, and setting $\sigma_0 = 0.5$, $v_{A0} = 0.58c$ and $\epsilon = 1$), remarkably close to the values found in Figure 3.11b.

Previous studies of imbalanced turbulence appear to exclude the possibility of generating net plasma motion along the background magnetic field, either through the assumption of reduced MHD, gyrokinetics, or force-free description [27, 28, 124]. As such, this work presents, to our knowledge, the first numerical demonstration and investigation of net flow due to imbalanced turbulence. Such a large net motion of the plasma may have implications for driving accretion disk winds, particularly if the wind comprises mostly nonthermal particles.

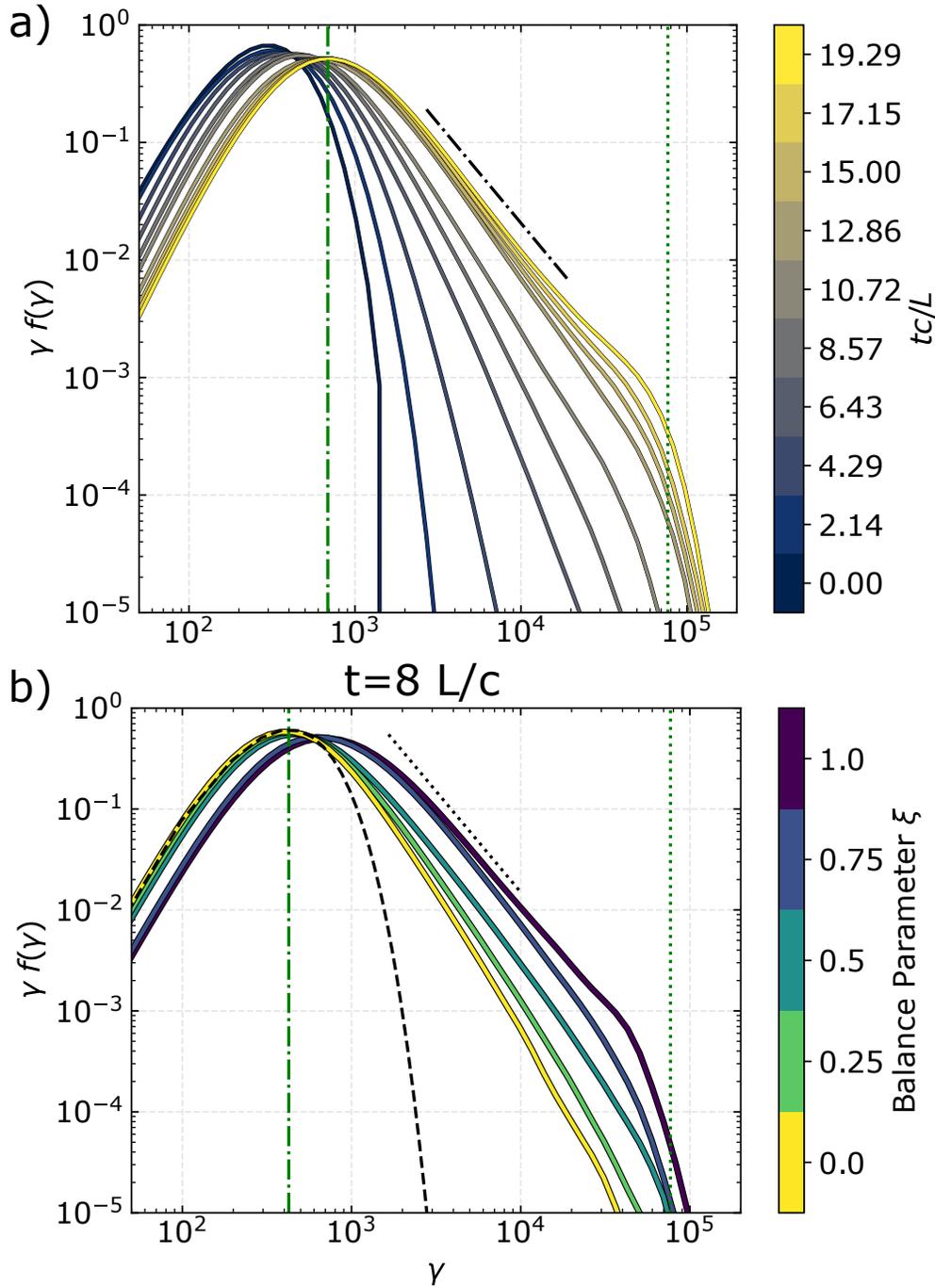


Figure 3.12: Particle acceleration occurs for all values of balance parameter. a) The distribution function of the most imbalanced case ($\xi = 0.0$) becomes shallower in time from an initial Maxwell-Jüttner distribution (purple) to a Maxwell-Jüttner distribution plus a hard power-law component at later times (yellow). b) Spectra taken at the same time $t = 8.0 L/c$ for different balance parameters show different peak energies but similar power-law components. The dashed black line shows a Maxwell-Jüttner fit to the $\xi = 0.0$ case. The vertical green dash-dot line shows the mean Lorentz factor $\langle \gamma \rangle$ extracted from this fit. The vertical green dotted line shows the maximum energy γ_{\max} . The dash-dot black line in panel (a) shows the power law $\gamma^{-2.7}$, while the dotted black line in panel (b) shows γ^{-3} . Colors are the same as in Fig. 3.3.

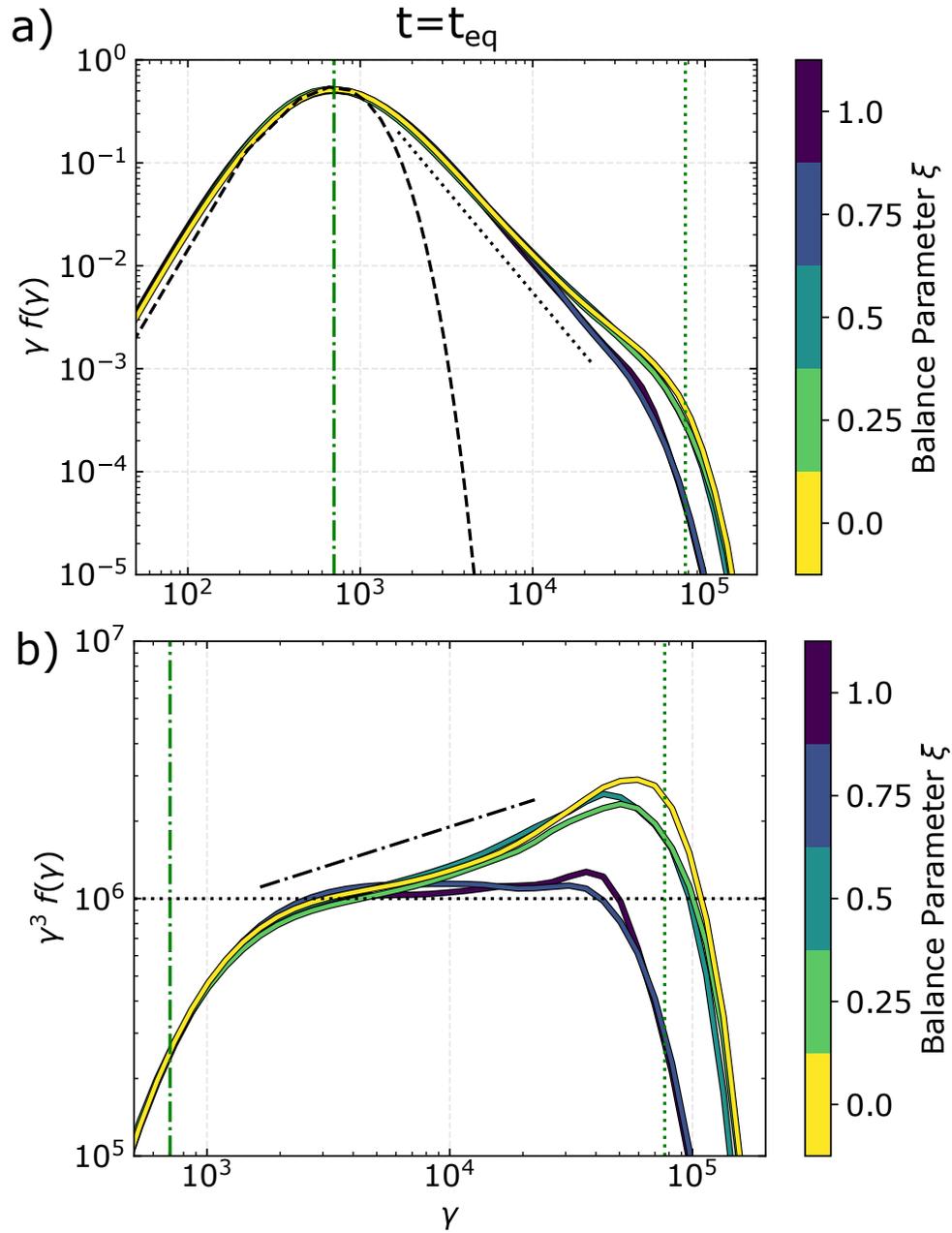


Figure 3.13: At equivalent times, the power laws of imbalanced turbulence are slightly flatter/harder than balanced turbulence. a) The particle energy spectra at equivalent times (see Table 3.3) show a similar mean energy and similar power laws until $\gamma \sim 10^4$. The dashed black line shows a Maxwell-Jüttner fit to the $\xi = 0.0$ case. b) Compensating by γ^3 reveals that more imbalanced turbulence ($\xi = 0.0, 0.25$, and 0.5) has flatter power laws than the more balanced turbulence with $\xi = 0.75$ or 1.0 . The vertical green dash-dot line shows the mean Lorentz factor $\langle \gamma \rangle$ extracted from the Maxwell-Jüttner fit. The vertical green dotted line shows the maximum energy γ_{\max} . The black dash-dot line in panel (b) shows the spectrum compensated to γ^3 , while the dotted black line shows the power law $\gamma^{-3+\beta}$ (a constant). Colors are the same as in Fig. 3.3.

3.3.4 Nonthermal Particle Acceleration

Nonthermal particle acceleration can explain high-energy flares and power laws seen in spectra of various astrophysical systems. Studying NTPA self-consistently requires PIC simulations. Recent PIC simulations of turbulence [86, 115, 116, 128] have successfully produced nonthermal particle populations that result in power-law spectra. Similar results have been produced in PIC simulations of kink-unstable jets [140, 141] and the magnetorotational instability in accretion discs [142–144], where turbulence may play a fundamental role in the particle acceleration. Because power-law spectra are observed in systems with asymmetric energy injection, it is important to understand how imbalance affects particle acceleration.

We find that imbalanced turbulence can accelerate a significant portion of the plasma’s particles to suprathermal energies (Fig. 3.12), much like balanced turbulence at similar magnetizations $\sigma \sim 1$. Even the most imbalanced case $\xi = 0.0$ shows the development of a high-energy power-law tail, which hardens and reaches an asymptotic slope after about $12 L/c$ (Fig. 3.12a). At late times $t \gtrsim 12 L/c$, the simulation domain’s boundary conditions limit the maximum attainable Lorentz factor to $\gamma_{\max} = LeB_0/m_e c^2$, resulting in the “pile-up” of high-energy particles at γ_{\max} , followed by a sharp cutoff rather than the continuation of the power law to even higher energies [114]. Visually, the nonthermal distribution matches the power-law scaling $f(\gamma) \propto \gamma^{-2.7}$ between $\langle \gamma \rangle$ and γ_{\max} , shown by the dash-dot line in Fig. 3.12a. Comparing turbulence with different balance parameters, we see that more balanced turbulence heats the plasma more quickly and forms a power-law spectrum faster than imbalanced turbulence (Fig. 3.12b). For particle spectra taken at $t = 8.0 L/c$, the simulation of balanced turbulence $\xi = 1.0$ has already heated to a peak Lorentz factor of about 700 and is experiencing pile-up, as shown by the spectrum’s break around $\gamma \sim 5 \times 10^4$. In contrast, the most imbalanced case $\xi = 0.0$ still has a peak Lorentz factor of around 400 and its power-law index has not yet reached an asymptotic value (Fig. 3.12b).

Because of the different rates at which energy is injected into simulations with different balance parameters, it may be more meaningful to compare particle spectra not at the same fixed absolute time, but rather at “equivalent times” when a fixed amount of energy has been injected. For definiteness, we take this time to coincide with the end ($t = 20 L/c$) of the simulation for the most imbalanced case, $\xi = 0$,

Table 3.3: Table of “equivalent” times t_{eq} in L/c where the same amount of energy density $10.3 B_0^2/8\pi$ has been injected. Values of the nonthermal particle and energy fraction at t_{eq} , shown in the second and third rows, are discussed in Section 3.3.4. Equivalent times are labelled in Fig. 3.2 as red \times 's.

ξ	0.0	0.25	0.5	0.75	1.0
t_{eq}	20.0	17.4	13.1	9.4	8.0
$N_{\text{nonthermal}}/N_{\text{total}}$	0.20	0.19	0.20	0.21	0.21
$\langle \mathcal{E}_{\text{nonthermal}} \rangle / \langle \mathcal{E}_{\text{pl}} \rangle$	0.54	0.53	0.55	0.54	0.54

corresponding to an injected energy of $10.3 B_0^2/8\pi$. The equivalent times for the simulations vary from $8 L/c$ for the balanced case $\xi = 1.0$ to $20 L/c$ for the most imbalanced case $\xi = 0.0$ (Table 3.3). When compared at these equivalent times, particle spectra for different ξ essentially collapse to a single universal curve (Fig. 3.13a). In particular, the peak Lorentz factors and the power-law tails up to $\gamma \approx 10^4$ become nearly indistinguishable. This similarity suggests that NTPA operates similarly in balanced and imbalanced turbulence when considered on appropriate timescales. In particular, the nonthermal segments of the distribution function match the power-law scaling $f(\gamma) \propto \gamma^{-3}$ (dotted line) for all values of balance parameter, suggesting that imbalanced and balanced turbulence accelerate particles with the same asymptotic spectra. Finer differences appear at higher energies B when the spectra are compensated by γ^3 ; whereas the balanced cases $\xi = 1.0$ and 0.75 closely follow this γ^{-3} scaling in the interval $200 \lesssim \gamma \lesssim 4 \times 10^4$, the more imbalanced cases are never quite flat and appear to more closely follow the scaling $\gamma^{-2.7}$ (dash-dot line), as shown in Fig. 3.13b. This power-law index of -2.7 matches the power-law index of balanced turbulence particle spectra in smaller box sizes [114], suggesting that pile-up contaminates the spectra. Because the equivalent times for the simulations of imbalanced turbulence are much longer, the high-energy pile-up could be due to a small sub-population of particles whose stochastic scattering events pushed them to higher energies. Larger simulation domains are needed to determine what influence the high-energy particle pile-up could have on the particle spectra at lower energies.

The partition of the plasma particles' energy $\langle \mathcal{E}_{\text{pl}} \rangle(t)$ into thermal and nonthermal components further demonstrates that nonthermal particles are energetically important in the system. The fraction of nonthermal particles is calculated by subtracting the Maxwell-Jüttner distribution that best fits the total, box-averaged particle distribution up to the peak Lorentz factor from the total distribution function. This fraction reaches

20% of the total number of particles at $20 L/c$ for the $\xi = 0$ case (Fig. 3.14a). The balanced case's fraction of nonthermal particles is larger, reaching 25% of the total number of particles by the same time. At $20 L/c$, the energy in these particles comprises 55% of the total plasma energy in the imbalanced case, as compared to 65% for the balanced case (Fig. 3.14b). At equivalent times, the nonthermal fractions of particles and energy do not vary more than 2% across balance parameters (Table 3.3), suggesting that particle acceleration by imbalanced turbulence is just as efficient as acceleration by balanced turbulence. For comparison, the fractions for the balanced simulations are slightly smaller than those found for similar simulations of electron-ion plasmas in the relativistically-hot limit [89].

Though highly idealized, quasi-linear theory can explain many aspects of turbulent NTPA. In particular, the treatment of NTPA as a diffusion in momentum space [126, 127] has been justified by measurements of the momentum diffusion coefficient in PIC simulations of balanced turbulence [116, 145] and improved by considering resonance broadening [146]. The original models suggest that the diffusion coefficient scales as $1 - \tilde{H}_c^2$, which is supported by test-particle simulations of imbalanced MHD turbulence when parallel acceleration is negligible [147]. Our simulations show that imbalance increases the acceleration timescale for NTPA, which is broadly consistent with a decrease in the diffusion coefficient. It is not clear why this increased acceleration timescale does not affect the power-law index.

3.4 Conclusions

In this study we investigate, for the first time, imbalanced kinetic turbulence in a collisionless, magnetized, relativistically hot plasma. Using 3D PIC simulations, we simulate a pair plasma driven by large-scale external currents, creating Alfvén waves propagating parallel and anti-parallel to the background magnetic field with different amplitudes. We demonstrate the formation of a turbulent cascade with a similar power-law index for all values of the balance parameter covered by the simulations (Section 3.3.1). We find that the energy injected into the plasma by the external driving is not only converted into internal energy through small-scale dissipative processes (Section 3.3.2), but also drives net bulk motion of the plasma (Section 3.3.3). This efficient transfer of momentum to the plasma appears as a relativistic effect, resulting in a net plasma velocity $\Gamma_{\text{net}} v_{\text{net}} \sim \delta \sigma v_{A0}$. We also find efficient particle acceleration over two decades of

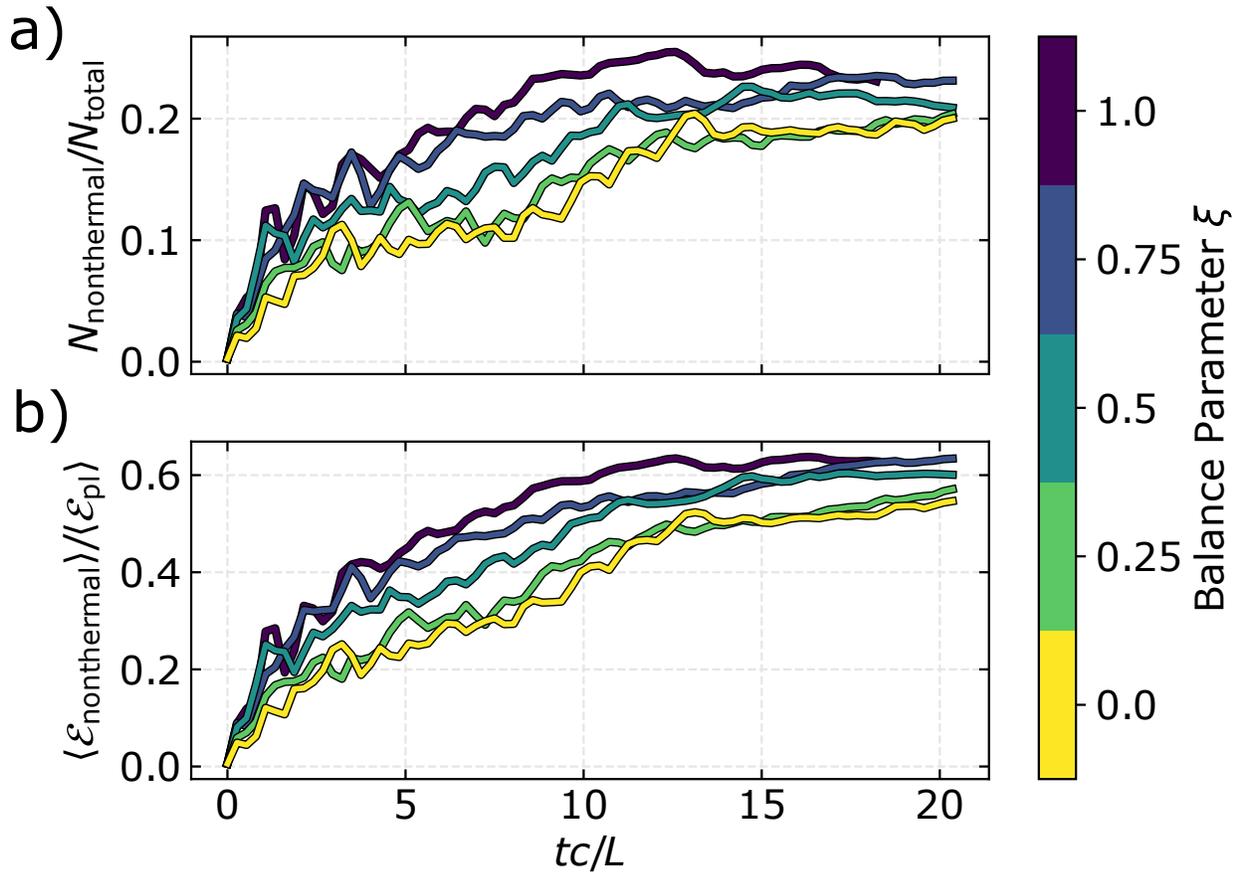


Figure 3.14: The partition of plasma energy $\langle \mathcal{E}_{\text{pl}} \rangle$ into thermal and nonthermal components shows a moderate increase with the balance parameter at any given time. Both the fraction of particles with nonthermal energies (a) and the fraction of total plasma energy density $\langle \mathcal{E}_{\text{pl}} \rangle$ contained in such particles (b) are calculated by fitting a thermal Maxwell-Jüttner function to the low- and medium-energy particle distribution at each time and subtracting the fit from the total particle distribution. Colors are the same as in Fig. 3.3.

particle Lorentz factor even for our most imbalanced turbulence (Section 3.3.4).

Our results on imbalanced turbulence should apply to high-energy astrophysical systems with asymmetric energy injection, such as accretion disk coronae, relativistic jets, and pulsar wind nebulae. We find that NTPA remains efficient in imbalanced turbulence, meaning that particle acceleration models developed for balanced turbulence are still applicable to astrophysical systems with asymmetric energy injection. In addition, our finding that the momentum from the driven Alfvén waves efficiently transfers to the plasma itself constitutes a new mechanism for propelling winds from, for instance, the surface of a turbulent accretion disk. This efficient momentum transfer could also amplify existing astrophysical outflows.

This work represents an important step in studying global properties of imbalanced turbulence in collisionless plasmas. It demonstrates a method for driving imbalanced turbulence in PIC simulations and develops diagnostics to study the unique aspects of imbalanced turbulence, including net motion of the plasma. Our study has revealed a number of factors that could influence the development of imbalanced turbulence and should be further explored: the driving mechanism, the amplitude of magnetic field fluctuations, and the plasma magnetization, to name a few. Our main findings of efficient NTPA and efficient momentum transfer merit further investigation: how are the Fokker-Planck momentum diffusion and advection coefficients for NTPA modified by imbalance? How does the momentum transfer manifest in a more realistic system with density gradients? Thus far, our finding of a net flow is a momentum-transfer mechanism, not a wind-launching mechanism. More work is needed to determine how the transfer efficiency changes with σ and whether the wind comprises the thermal bulk of particles or nonthermal particles. Simulations of imbalanced turbulence in nonrelativistic, semi-relativistic, and trans-relativistic electron-ion plasmas, particularly relevant to accretion flows, will also be important for understanding the fraction of energy that heats electrons. Understanding these aspects of imbalanced turbulence will aid in modeling astrophysical systems with asymmetric energy injection, such as accretion disk coronae, relativistic jets, and pulsar wind nebulae.

Acknowledgements

AMH acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1650115. VZ acknowledges support from NASA Hubble Fellowship grant HST-HF2-51426.001-A. This work was also supported by NASA ATP grants NNX17AK57 and 80NSSC20K0545, and NSF Grants AST-1806084 and AST-1903335.

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 [148]. We used the XSEDE resources Stampede2 and Ranch at the Texas Advanced Computing Center (TACC) through allocations PHY-140041 and PHY-160032.

Data Availability

The simulation data underlying this article were generated at the XSEDE/TACC Stampede2 super-computer and are archived at the TACC/Ranch storage facility. As long as the data remain in the archive, they will be shared on reasonable request to the corresponding author.

3.A Dependence on Domain Size

If the ratio $L/2\pi\rho_{e0}$ is small, the small separation between the characteristic scale of kinetic effects and the system size scale could influence the results presented in Section 3.3. In this appendix, we examine the box-size dependence of representative quantities for extremal values of the balance parameter $\xi = 0.0$ and $\xi = 1.0$ and for $L/2\pi\rho_{e0} \in \{27.1, 40.7, 54.3, 81.5\}$, corresponding to $N \in \{256, 384, 512, 768\}$. These results also include three very large simulations of balanced turbulence from Ref. [114] with $L/2\pi\rho_{e0} \in \{81.5, 108.7, 164\}$ ($N = 768, 1024, \text{ and } 1536$) that are otherwise identical to the other simulations of balanced turbulence. The time-averaging window has been changed from $10 < tc/L < 20$ to $5 < tc/L < 14$, because $14 L/c$ is the latest time included in all simulations. Here we focus on the convergence of energetic quantities; for convergence of the balanced turbulence's particle energy spectra with system size, see Ref. [114].

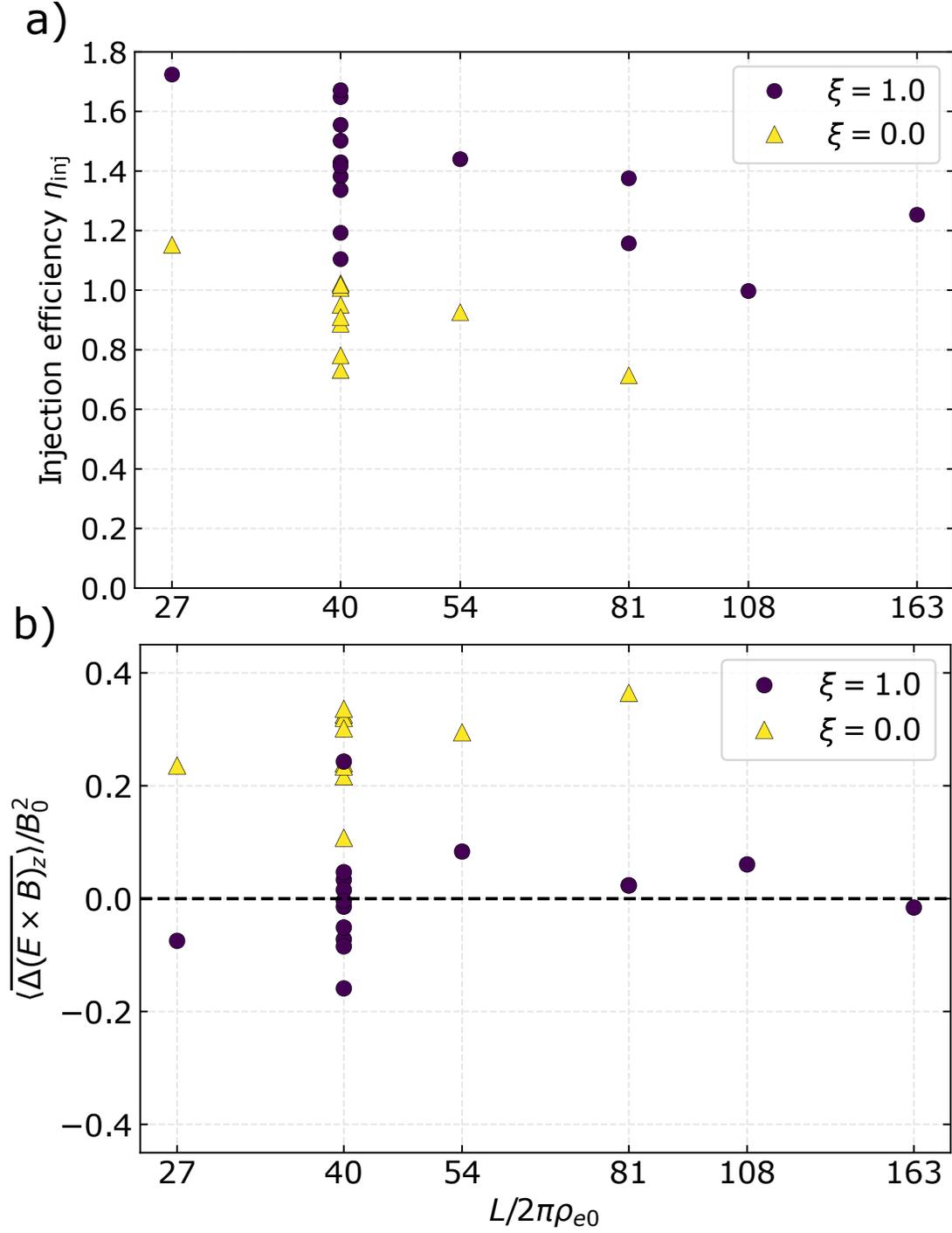


Figure 3.15: The injection efficiency and Poynting flux along the background magnetic field depend weakly on simulation domain size within statistical variation. When averaged from $t = 5 - 14 L/c$, the injection efficiency (a) and Poynting flux along the background magnetic field (b) are shown as a function of $L/2\pi\rho_{e0}$. Both the balanced ($\xi = 1.0$; purple circles) and most imbalanced values ($\xi = 0$; yellow triangles) are mostly within statistical variation of the $L/2\pi\rho_{e0} = 40.7$ sample of 8 random seeds. The black dashed line indicates zero.

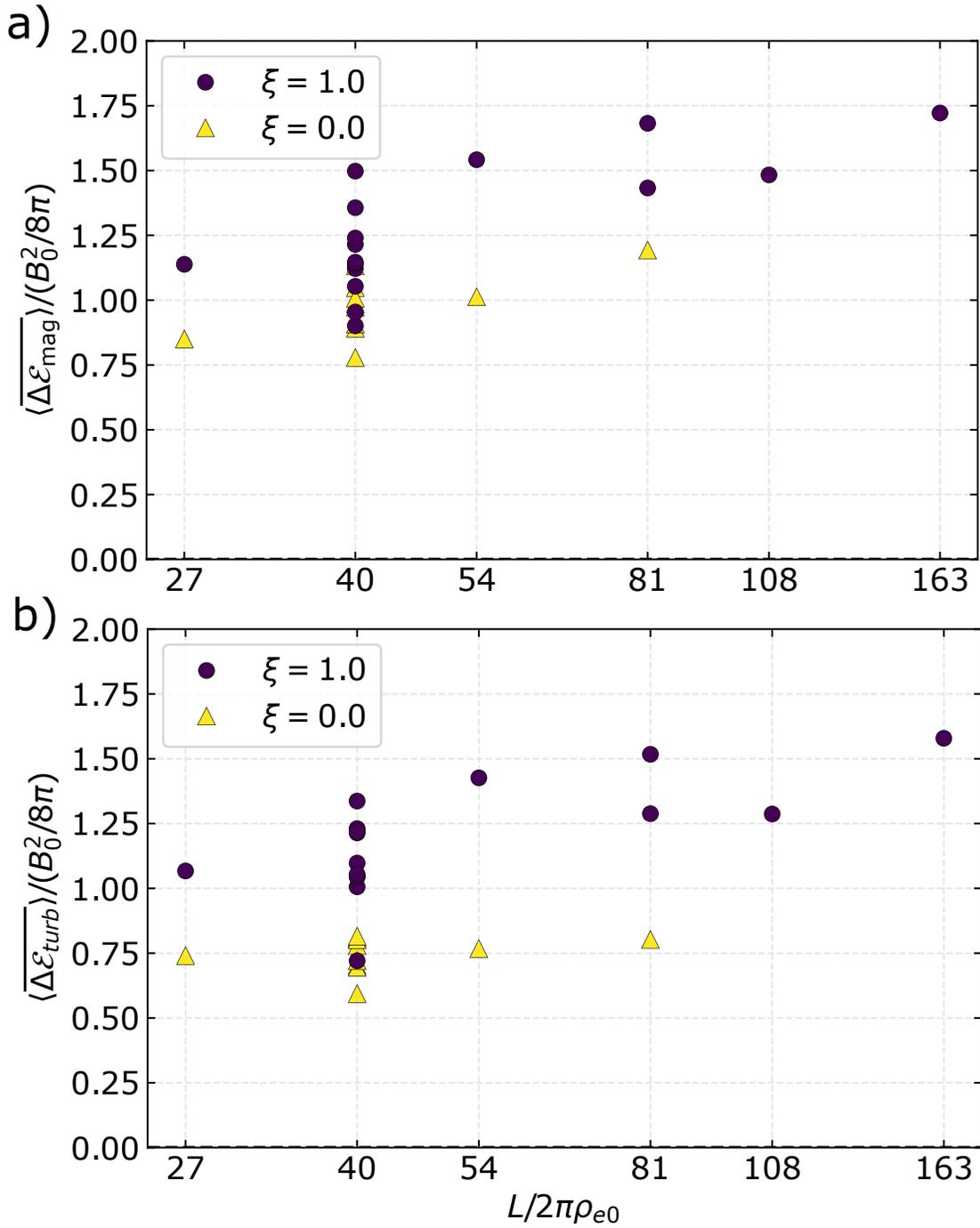


Figure 3.16: The magnetic and turbulent kinetic densities are weakly dependent on simulation domain size. When averaged from $t = 5 - 14 L/c$, the magnetic energy density (a) and turbulent kinetic energy density (b) are shown as a function of $L/2\pi\rho_{e0}$. Both the balanced ($\xi = 1.0$; purple circles) and most imbalanced ($\xi = 0$; yellow triangles) values are mostly within statistical variation of the $L/2\pi\rho_{e0} = 40.7$ sample of 8 random seeds.

We find that the injection efficiency η_{inj} depends weakly on simulation domain size for both balanced and imbalanced turbulence (Fig. 3.15a). The simulations with $\xi = 1.0$ and $L/2\pi\rho_{e0} = 81.5$ and 164 domains have slightly lower injection efficiencies than those for the smallest ($L/2\pi\rho_{e0} = 40.7$) domains. Kinetic damping of large-scale fluctuations, which would drain energy faster than turbulence alone, may explain the larger η_{inj} for smaller domain sizes. For $\xi = 0$, the $L/2\pi\rho_{e0} = 54.3$ simulation's η_{inj} is within the statistical spread of the $L/2\pi\rho_{e0} = 40.7$ simulations' injection efficiencies, whereas the $L/2\pi\rho_{e0} = 81.5$ injection efficiency is slightly below. The time-averaged Poynting flux shows a weak positive trend with increasing domain size (Fig. 3.15b). Though the $L/2\pi\rho_{e0} = 81.5$ domain size for the imbalanced $\xi = 0$ case has a value higher than the largest $L/2\pi\rho_{e0} = 40.7$ value, the difference is only about 0.05 ($cB_0^2/4\pi$) (about 15%), within two standard deviations of the statistical variation shown by the $L/2\pi\rho_{e0} = 40.7$ study.

The turbulent and magnetic energy densities show a slight dependence on simulation domain size (Fig. 3.16). Both quantities' values for $L/2\pi\rho_{e0} \gtrsim 54.3$ are consistently about 15% greater than the largest value of the statistical ensemble of $L/2\pi\rho_{e0} = 40.7$, $\xi = 1.0$ simulations. Though the imbalanced turbulence simulations do not appear to exhibit a trend in turbulent kinetic energy with box size (Fig. 3.16b), the $L/2\pi\rho_{e0} = 81.5$, $\xi = 0$ case's value for magnetic energy is noticeably higher (20%) than the $L/2\pi\rho_{e0} = 54.3$ value, which lies within the statistical spread of the $L/2\pi\rho_{e0} = 40.7$ ensemble study (Fig. 3.16a).

Chapter 4

Nonthermal Emission from the Plunging Region: a Model for the High-Energy Tail of Black Hole X-ray Binary Soft States

In this chapter, I build upon the idea that nonthermal electrons can be produced in an accretion disk corona. More specifically, I show that the compact region within the ISCO of a thin accretion disk (the “plunging region”) has conditions conducive to nonthermal particle acceleration. This work was motivated by GRMHD simulations that suggest the assumption of single-temperature MHD within ISCO no longer holds, and observations showing the physical compactness of the corona (Ch. 2.4.2). I use an analytic model of the plunging region to provide a background for a nonthermal electron distribution. I investigate whether the hierarchy of particle timescales allows for the existence of nonthermal particles, and whether GR effects would wipe out any radiation from within the ISCO. I use prescriptions from PIC simulations for particle acceleration to model the electron distribution function at each radius inside the plunging region. This model takes a step towards building in PIC prescriptions to global models of the accretion disk.

This chapter’s model for the corona agrees well with observations. I predict some trends with black hole spin and the inclination of the accretion disk with respect to the observer that could be tested against future observations. This model is an attractive way to connect the disk truncation model (Ch. 2.3.1) to the soft state of BHBs (Ch. 2.4.1). It provides a framework for locally linking kinetic physics and background global dynamics and can be improved upon to include better particle acceleration prescriptions as they emerge. This work was published in Ref. [5].

Abstract

X-ray binaries exhibit a soft spectral state comprising thermal blackbody emission at 1 keV and a power-law tail above 10 keV. Empirical models fit the high-energy power-law tail to radiation from a nonthermal electron distribution, but the physical location of the nonthermal electrons and the reason for their power-law index and high-energy cut-off are still largely unknown. Here, we propose that the nonthermal electrons originate from within the black hole’s innermost stable circular orbit (the “plunging region”). Using an analytic model for the plunging region dynamics and electron distribution function properties from particle-in-cell simulations, we outline a steady-state model that can reproduce the observed spectral features. In particular, our model reproduces photon indices of $\Gamma \gtrsim 2$ and power-law luminosities on the order of a few percent of the disk luminosity for strong magnetic fields, consistent with observations of the soft state. Because the emission originates so close to the black hole, we predict that the power-law luminosity should strongly depend on the system inclination angle and black hole spin. This model could be extended to the power-law tails observed above 400 keV in the hard state of X-ray binaries.

4.1 Introduction

Observations of accreting black hole X-ray binary systems (XRBs) show two main X-ray spectral states. The hard state has an inverted power-law photon spectrum $N(E) \propto E^{-\Gamma}$ with a spectral index $\Gamma \lesssim 2$ at 10 keV, peaking in the 100 – 200 keV range. The soft state has a thermal blackbody component peaking around 1 keV and a high-energy power-law tail component from 10 keV to 1 MeV. This hard X-ray to soft gamma-ray power law has a spectral index $\Gamma \gtrsim 2$ and no observed cut-off [48, 149, 150].

Models of the soft spectral state usually involve a thin accretion disk that produces the thermal emission and a hot, optically-thin gas that produces the high-energy power-law tail. The thin accretion disk [9, 49] models the soft state’s blackbody emission well. However, the origin of the power-law tail is less well-understood. Most models for the power-law tail assume a population of nonthermal electrons that are somehow accelerated to high energies, presumably by shocks or magnetic reconnection. The hybrid thermal/nonthermal electron distribution function fits the soft state spectrum up to 600 keV well [151, 152],

but the electron distribution parameters are determined by the best fit to observational data rather than an initial theory.

The spatial location of the hot, optically-thin gas and thus the nonthermal electrons is unknown. Hybrid electron distribution models usually assume that the nonthermal electrons originate from a jet or corona above a thin accretion disk extending down to the innermost stable circular orbit [ISCO; 150]. Alternatively, the nonthermal electrons could come from gas within the ISCO. In the “plunging region,” particle orbits transition from predominantly circular to predominantly radial. The ISCO serves as an artificial boundary condition in early thin disk models, which assume that the stress there goes to zero and no light is emitted from within the plunging region [9, 49]. Later thin disk models find that stress at the ISCO can lead to extra dissipation, which is often parameterized as an additional accretion efficiency $\Delta\epsilon$ [50, 51]. Magnetohydrodynamic (MHD) and general relativistic MHD (GRMHD) simulations have measured additional dissipation within the ISCO on the order of $\Delta\epsilon \gtrsim 0.2$ for weakly magnetized disks and up to twice the Novikov-Thorne efficiency for strongly magnetized disks [63, 153–156]. This extra dissipation could come from magnetic reconnection and turbulence occurring in the plunging region.

Extra dissipation from within the plunging region could cause the soft state’s high-energy power-law tail. In a purely thermal model, the gas temperature rises so rapidly towards the event horizon that the Wien peaks sum to a power law [157]. This model could also explain high-frequency quasi-periodic oscillations in the steep power law state [158]. These models assume that the electron and proton gas are in thermal equilibrium. However, shorter infall times in the plunging region can mean electrons and protons no longer have time to thermalize. Estimates of the electron-proton thermalization time in GRMHD simulations become longer than the infall time close to the black hole, suggesting that electrons and protons may decouple into a two-temperature plasma within the ISCO [157].

In this work, we present a model for the soft spectral state’s high-energy power-law tail using non-thermal electrons in the plunging region. We outline the parameter space where electron-proton decoupling should occur and demonstrate that the plunging region conditions are conducive to accelerating electrons to nonthermal energies (Section 4.2). We then propose a model for the hybrid thermal/nonthermal electron distribution in the plunging region based on particle acceleration prescriptions from particle-in-cell simula-

tions (Section 4.3). We show that our model can reproduce observational characteristics of a high-energy power-law tail for highly-magnetized parameters, recover the transition to thermal-dominated spectra at lower magnetizations, and predict trends with inclination angle and black hole spin that are consistent with current observations (Section 4.4). We end by discussing our model assumptions and its general applicability to XRBs (Section 4.5) and summarizing our findings (Section 4.6).

4.2 Dynamical Model for the Plunging Region

Within the ISCO, the accretion flow cannot be described as a thin disk. Particle orbits transition from mostly circular outside the ISCO to mostly radial inside the ISCO. The fast infall times in the plunging region mean that the usual assumptions of a collisional, single-temperature plasma need to be revisited. In this section, we will motivate our treatment of the plunging region as a two-temperature, highly-magnetized region with the potential to continuously accelerate particles to nonthermal energies.

We will use Ref. [51]’s plunging region model as a backdrop for our model. Ref. [51] provides radial profiles of the gas number density n_0 , four-velocity u^μ , and radial magnetic field strength B_r that agree well with numerical simulations [63]. The model assumes a one-dimensional, steady-state MHD inflow in the plunging region, parameterized by black hole spin a , mass M , and the $F_{\theta\phi}$ component of the Maxwell tensor, which reduces to $-r^2 B_r$ at the midplane in the nonrelativistic limit. Throughout, we assume an Eddington factor $\eta = 0.1 = L_{\text{Edd}}/(\dot{M}_{\text{Edd}}c^2)$, where L_{Edd} and \dot{M}_{Edd} are the Eddington luminosity and accretion rate, respectively. We avoid numerical artefacts due to the boundary condition $v^r = 0$ at the ISCO by starting our solutions at $r_{\text{ISCO}} - 0.1r_g$. We scale the model to physical units using a black hole mass of $10M_\odot$ and a density scale height $H/r = 1$, the latter choice motivated by Section 4.2.1’s observation that ions will not collide with electrons in the plunging region.

For our fiducial model, we use values $a = 0.95$ and $F_{\theta\phi} = 6.0$, which corresponds to an accretion efficiency $\Delta\epsilon = 1.04$. For all models, we assume that $\Delta\epsilon$ is uniformly distributed throughout the volume within the ISCO and that this heating comes from magnetic reconnection processes. Representative quantities of the fiducial background are plotted in Figure 4.1. For these parameters, the magnetic field has a strength on the order of 10^8 G. We assume a virial ion temperature $T_i = 2GMm_p/5k_B r$, which gives mildly relativistic

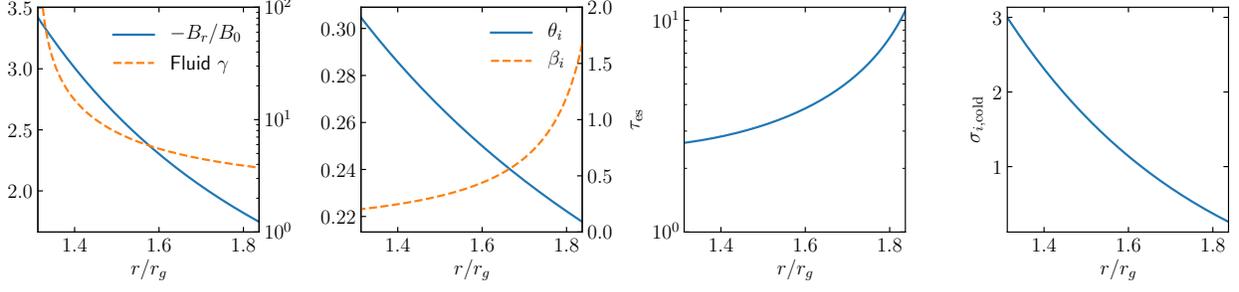


Figure 4.1: Properties of the plunging region calculated from the Ref. [51] background for fiducial parameters $a = 0.95$ and $\Delta\epsilon = 1.04$. From left to right: radial magnetic field strength B_r (left axis) and the fluid’s Lorentz factor (right axis), dimensionless ion temperature $\theta_i = k_B T_i / m_p c^2$ (left axis) and ratio of gas-to-magnetic pressure $\beta_i = 8\pi n_0 k_B T_i / B_r^2$ (right axis), electron scattering optical depth τ_{es} , and cold ion magnetization σ_i . Here, $B_0 = 10^8$ G. The plunging region extends from the ISCO at $1.94r_g$ to the event horizon at $1.31r_g$ in the midplane.

ions, and calculate the ion beta $\beta_i = 8\pi n_0 k_B T_i / B_r^2$. The electron scattering optical depth $\tau_{es} = n_0 \sigma_T H$ is always slightly greater than 1. The values of B_r , τ , and β_i vary by a factor of ~ 3 from the ISCO to the event horizon.

4.2.1 Development of a Two-temperature Plasma

Analytic models of a thin disk that permit non-zero stress at the ISCO predict a factor of ten increase in the gas temperature between the ISCO and the event horizon [50]. Single-temperature GRMHD simulations find a similar increase [157] due to the flow becoming effectively optically thin. The rapid rise of temperature means that the timescale for electrons and protons to reach thermal equilibrium increases as well. This electron-ion thermalization timescale t_{th}^{ei} is given by

$$\frac{dT_e}{dt} = \frac{T_i - T_e}{t_{th}^{ei}} \quad (4.1)$$

where $t_{th}^{ei} \equiv 1/\nu_{th}$. For hot protons heating electrons, the thermalization frequency from Ref. [102] reads:

$$\nu_{th} = \frac{m_e}{m_p} \frac{n_0 \sigma_T c \ln \Lambda}{K_2(1/\theta_e) K_2(1/\theta_i)} \quad (4.2)$$

$$\left[\frac{2(\theta_e + \theta_i)^2 + 1}{\theta_e + \theta_i} K_1 \left(\frac{\theta_e + \theta_i}{\theta_e \theta_i} \right) + 2K_0 \left(\frac{\theta_e + \theta_i}{\theta_e \theta_i} \right) \right] s^{-1} \quad (4.3)$$

$$= 3.6 \times 10^{-37} \frac{(m_e m_p)^{1/2} n_0 \ln \Lambda}{(m_p T_e + m_e T_i)^{3/2}} s^{-1}, \quad (4.4)$$

where $\theta_e = k_B T_e / m_e c^2$ and $\theta_i = k_B T_i / m_p c^2$, m_e and m_p are electron and proton masses, respectively. The last line takes the nonrelativistic limit assuming $\theta_e \ll 1$ and $\theta_i \ll 1$ [159]. Here, $K_0(x)$, $K_1(x)$, and $K_2(x)$ are modified Bessel functions of the second kind. Throughout, we set the Coulomb logarithm $\ln \Lambda = 10$ and employ cgs units.

In the plunging region, the thermalization timescale can become longer than the accretion timescale. Ref. [157] finds that a weakly magnetized flow cannot equilibrate within $3 r_g$ for black hole spin $a = 0.7$. Therefore, the assumption that electrons and protons couple efficiently no longer applies within the ISCO. The inefficient Coulomb coupling means that protons cannot cool, and should heat due to viscous and magnetic dissipation. Such increased temperature and thus increased ion pressure support will puff up the plunging region beyond a thin disk. The formation of a thick inner disk and a thin outer disk was recently simulated for the first time [80]. For those model parameters, the thin outer disk with $H/r \sim 10^{-2}$ expanded to $H/r \sim 0.3$ within the ISCO.

This decoupling motivates our subsequent treatment of the electrons and protons in the plunging region as completely independent, with different temperatures, as well as our assumption that $H/r \sim 1$. Indeed, Figure 4.2 shows that in our model, the thermalization time (black solid line) is longer than the infall time (dashed horizontal line) for a wide range of accretion efficiencies. The same hierarchy holds for all black hole spins discussed in this work.

4.2.2 Nonthermal Particle Acceleration

Particle-in-cell (PIC) simulations of magnetic reconnection have shown that efficient particle acceleration in an electron-ion plasma occurs for ion magnetizations $\sigma_i \gtrsim 1$ [68]. Using background profiles for magnetic field and gas density from Ref. [51], we will demonstrate that efficient particle acceleration is possible inside the ISCO. We calculate the ion magnetization

$$\sigma_i = \frac{B_0^2}{4\pi n_0 m_p c^2} \quad (4.5)$$

using the definition from Ref. [68]. We set $B_0 = B_r$ and $n_0 = \rho/m_p$, with the radial magnetic field $B_r(r)$ and mass density $\rho(r)$ obtained from Ref. [51]. For fiducial model parameters, the ion magnetization ranges

from 0.2 to 3.0 (Figure 4.1). In PIC simulations of magnetic reconnection with similar ion magnetizations, Ref. [68] finds that electrons are accelerated into power-law distribution with high-energy cut-off $\gamma_c \sim 4\sigma_e$ (rising slowly with σ_i) and power-law index p prescribed by

$$p = 1.9 + 0.7/\sqrt{\sigma_i}. \quad (4.6)$$

Ref. [68] fixes $\theta_i = \sigma_i/200$ (i.e. $\beta_i = 4\theta_i/\sigma_i \rightarrow 0.02$), a factor of 10 to 100 times colder than the equipartition ion temperature calculated above. PIC studies of electron-ion turbulence (rather than magnetic reconnection) that set $\beta_i = 2/3$ and scanned $\theta_i \in [1/256, 10]$ with $\sigma_i \in [1/256, 10]$ have found broadly similar power-law indices [89]. The turbulent simulations in general find more efficient particle acceleration, measuring $p \approx 2.8$ for $\theta_i = 1/16$ and $\sigma_i = 0.2$ when Equation 4.6 would predict $p = 3.5$.

With ion magnetizations close to unity, the plunging region can plausibly accelerate electrons to nonthermal energies via turbulence or reconnection. This acceleration could happen in a two-stage process, with sharp increases in particle Lorentz factor when the particle crosses a current sheet and slower, second-order Fermi-like acceleration that continues to increase particle energy over time [116]. For a magnetically-dominated pair plasma, the current sheet acceleration time is on the order of $t_{\text{acc}} \sim (300/\sigma)\omega_{pe}^{-1}$ [116], where $\omega_{pe} = \sqrt{4\pi e^2 n_0/m_e}$ is the electron plasma frequency. For the fiducial parameters $\Delta\epsilon = 1.04$ and $a = 0.95$, the Ref. [51] model yields $100\omega_{pe}^{-1} \sim 10^{-11}$ s, much shorter than infall times on the order of 10^{-4} s for a ten solar-mass black hole and electron-ion thermalization times (Figure 4.2 orange line). Therefore, we argue that the plunging region can accelerate electrons locally, forming a power-law distribution that will not have time to be thermalized by protons.

4.2.3 Electron-electron Collisions

A pure power-law distribution of electrons will interact via Coulomb collisions, eventually forming a pure thermal distribution if allowed to evolve freely [160]. Lower-energy electrons thermalize first, forming a hybrid thermal/nonthermal population at intermediate evolution times. If nonthermal electrons are continuously injected, as Section 4.2.2 showed is possible in the plunging region, the electron distribution will reach some equilibrium of nonthermal and thermal populations. We model this hybrid distribution as a

thermal Maxwell-Jüttner distribution up to a transition Lorentz factor γ_1 , where the distribution becomes a power law. A nonthermal electron with Lorentz factor γ will interact with the thermal bulk on the energy loss timescale $t_{\text{coll}}^{ee}(\gamma) \equiv 1/\nu_{\epsilon}^{ee}(\gamma)$. The energy loss rate is given by

$$\nu_{\epsilon}^{ee}(\gamma) = [\psi(x^{ee}) - \psi'(x^{ee})] \frac{16\sqrt{\pi}e^4 \ln \Lambda n_0}{m_e^2 c^3 \beta^3}, \quad (4.7)$$

where $\beta = \sqrt{1 - 1/\gamma^2}$ and $\psi(x)$ is the lower incomplete Gamma function [159]. The kinetic ratio $x^{ee} = \beta^2/(2\theta_e)$. For an electron temperature of 10^9 K, a particle with a Lorentz factor of 2.0 has an energy loss timescale of 10^{-6} s, about 0.002 times the infall timescale (Figure 4.2 purple). We therefore cannot neglect electron-electron collisions, especially at lower electron energies.

Though nonthermal electrons will also collide with thermal protons over the corresponding electron-proton energy loss timescale, this timescale will be larger than the thermalization timescale discussed in Section 4.2.1. We will therefore ignore interactions between nonthermal electrons and thermal ions.

4.2.4 Importance of Radiative Cooling

High-energy electrons cool rapidly due to either synchrotron or inverse Compton cooling. In our model, the cooling time $t_{\text{cool}}(\gamma)$ is set by either synchrotron or inverse Compton cooling, depending on the ratio U_B/U_{ph} , where $U_B = B_r^2/8\pi$ is the magnetic energy density and $U_{ph} = L/(4\pi c r_{\text{ISCO}}^2)$ is the energy density in the ambient photons, assumed to come from the disk such that $L = 0.1 L_{\text{Edd}}$, where $L_{\text{Edd}} = 4\pi G M m_p c / \sigma_T = 1.26 \times 10^{39}$ erg/s. If $U_B/U_{ph} > 1$, then the cooling time is set to the time it takes for an electron to lose half its energy due to synchrotron radiation [57]:

$$t_{\text{sync}}(\gamma) = \frac{5.1 \times 10^8}{\gamma B_r^2}. \quad (4.8)$$

In using the above equation, we assume that the particle is moving in gyro-orbits around the dominant magnetic field B_r , with a pitch-angle $\alpha = \pi/2$. If on the other hand $U_B/U_{ph} < 1$, the cooling time is set to the inverse Compton cooling time, assumed to take the form

$$t_{\text{IC}}(\gamma) = \frac{U_B}{U_{ph}} t_{\text{sync}}(\gamma). \quad (4.9)$$

Notice that Equation 4.8 and 4.9 are valid only for ultra-relativistic particles. As such, we ignore cooling for electrons below a minimum Lorentz factor $\gamma_{\min} = 2$, since below this Lorentz factor the cooling rate decreases rapidly.

The cooling times decrease for larger Lorentz factors, with values at the minimum Lorentz factor two orders of magnitude smaller than the electron-electron collision time, and five orders of magnitude smaller than the infall timescale (Figure 4.2 magenta). The electrons in the plunging region therefore have time to radiate before disappearing beyond the event horizon.

We show the relevant timescales of the plunging region in Figure 4.2 for models with different black hole spin a and accretion efficiencies $\Delta\epsilon$ resulting from different magnetic field strengths. At a radius of $r = 1.58 r_g$, all models satisfy the ordering

$$t_{\text{accel}} \ll t_{\text{cool}}(\gamma) \ll t_{\text{coll}}^{ee}(\gamma) \ll t_{\text{infall}} \lesssim t_{\text{therm}}^{ei} \quad (4.10)$$

for $\gamma \geq \gamma_1$. We will use this ordering to motivate our model for the plunging region's hybrid thermal/nonthermal electron distribution.

4.3 Multizone Equilibrium Model

In this section, we present a model for the electron distribution in the plunging region, motivated by the properties outlined in Section 4.2. We will assume that nonthermal electrons are continuously injected by local processes such as magnetic reconnection or turbulence, which then partially thermalize via electron-electron collisions. We further assume that the hybrid thermal/nonthermal electron distribution reaches a steady state at each radius, matching the losses due to radiation with the heating due to magnetic dissipation. Below, we describe the assumptions that allow us to fully constrain this hybrid electron distribution's form.

4.3.1 Initial Electron Distribution

Because the advection timescales are much longer than the particle acceleration timescales (Figure 4.2), we assume that the initial (pre-thermalization) electron population at each radius is a pure power-law distribution: $f_{\text{PL},0}(\gamma) = A_0\gamma^{-p}$ for $1 \leq \gamma \leq \gamma_2$, where γ_2 is a high-energy cut-off. The high-energy cut-off

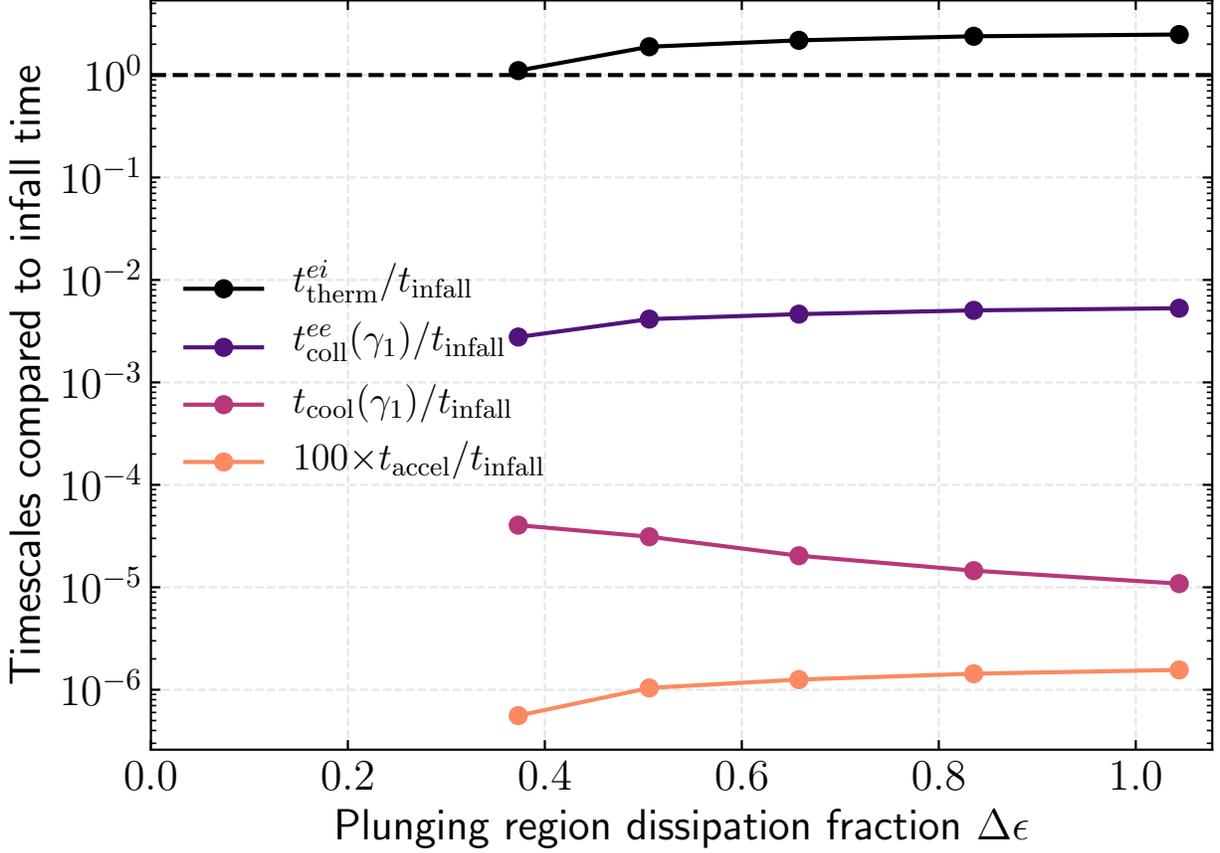


Figure 4.2: Timescales in the plunging region show that all physical processes except for electron-ion collisions are fast compared to the infall time for a wide range of accretion efficiencies $\Delta\epsilon$, consistent with the hierarchy $t_{\text{accel}} \ll t_{\text{cool}}(\gamma) \ll t_{\text{coll}}^{ee}(\gamma) \ll t_{\text{infall}} \lesssim t_{\text{therm}}^{ei}$. Timescales are shown at a radius $r = 1.58r_g$ for black hole spin $a = 0.95$. These timescales are calculated for thermal ion and hybrid electron distributions determined by the steady-state model described in Section 4.3.

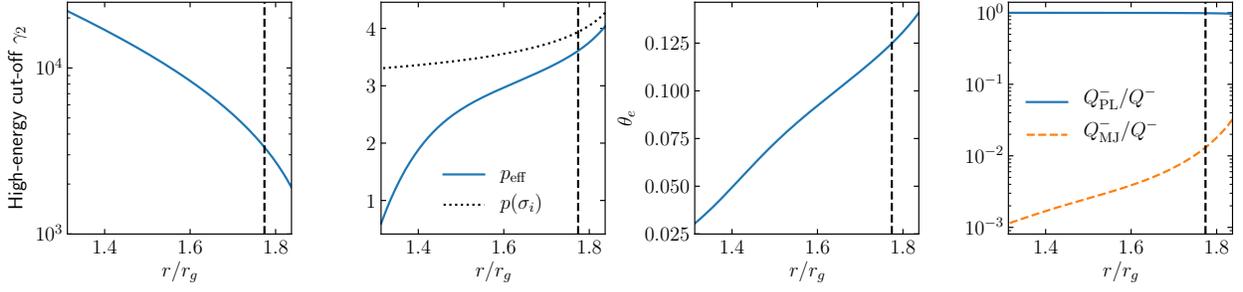


Figure 4.3: Radial profiles of the electron distribution function model properties for fiducial parameters $a = 0.95$ and $\Delta\epsilon = 1.04$. From left to right: the high-energy Lorentz factor cut-off of the power law γ_2 , the electron power-law index, the thermal electron temperature $\theta_e = k_B T_e / m_e c^2$, and the fraction of the total cooling rate Q_- from the power-law and thermal component. Vertical dashed line shows the location of the half-light radius.

at each radius is set to

$$\gamma_2(r) = \gamma_c(r) = 4\sigma_e(r), \quad (4.11)$$

where $\sigma_e = B_r^2(r)/(4\pi n_0 m_e c^2)$ is the (cold) electron magnetization [68]. The power-law index p is also set by Ref. [68]’s prescription (Equation 4.6), increased by 1 to account for cooling [161]. Requiring the distribution to be normalized to n_0 sets the constant A_0 . All together, the initial distribution reads:

$$f_{\text{PL},0}(\gamma) = \frac{n_0(p-1)}{1-\gamma_2^{1-p}} \gamma^{-p} \quad 1 \leq \gamma \leq \gamma_2. \quad (4.12)$$

Electron-electron Coulomb collisions will partially thermalize this initial power-law distribution. Radial profiles for the high-energy cut-off γ_2 and the power-law index p in the fiducial model are given in Figure 4.3.

4.3.2 Steady-state Electron Distribution

Some of the injected power-law electrons will interact with each other via Coulomb collisions, creating a thermal distribution below some Lorentz factor γ_1 (Section 4.2.3). The resulting hybrid thermal/nonthermal electron distribution continuously gains more nonthermal electrons due to the injection process and loses nonthermal electrons to the thermal bulk, radiating energy. We approximate these rapid processes with a steady-state electron distribution comprising a Maxwell-Jüttner thermal distribution f_{MJ} and a power-law distribution $f_{\text{PL}}(\gamma)$ for $\gamma_1 \leq \gamma \leq \gamma_2$.

The thermal electron distribution f_{MJ} is fully described by two variables: its normalization and its temperature T_e . We fix the normalization of the thermal distribution to the number density n_0 , assuming that the number of nonthermal electrons is negligible compared to the thermal electrons. Explicitly, the thermal distribution reads

$$f_{\text{MJ}}(\gamma, \theta_e, n_0) = n_0 \frac{\gamma^2 \beta}{\theta_e K_2(1/\theta_e)} e^{-\gamma/\theta_e}, \quad (4.13)$$

where K_2 is the modified Bessel function of the second kind. The electron temperature $T_e = \theta_e m_e c^2 / k_B$ is set by the model assumptions outlined below.

The nonthermal electrons are described as a power-law between a minimum Lorentz factor γ_1 and a high-energy cut-off γ_2 , with a power-law index p_{eff} and a normalization A_{PL} :

$$f_{\text{PL}}(\gamma) = A_{\text{PL}} \gamma^{-p_{\text{eff}}} \quad \gamma_1 \leq \gamma \leq \gamma_2. \quad (4.14)$$

We will now outline how our model determines the variables γ_1 , A_{PL} , p_{eff} , and T_e .

4.3.2.1 Determining γ_1

The minimum Lorentz factor γ_1 of the nonthermal distribution is set at the energy where collisions and cooling balance. At γ_1 , the energy loss timescale t_{coll}^{ee} (Equation 4.7) equals the cooling time t_{cool} :

$$t_{\text{coll}}^{ee}(\gamma_1) = t_{\text{cool}}(\gamma_1). \quad (4.15)$$

Here, the cooling time t_{cool} is set to the minimum of the synchrotron and inverse Compton cooling times (Equations 4.8 and 4.9). We enforce the lower bound $\gamma_{1,\text{min}} = \gamma_{\text{min}} = 2.0$ to model the inefficiency of cooling at low energies. We also require that $\gamma_1 > \gamma_{\text{threshold}}(T_e)$, where the threshold Lorentz factor $\gamma_{\text{threshold}}$ is where the energy loss frequency (Equation 4.7) is zero for a given temperature. A power-law electron with a Lorentz factor below $\gamma_{\text{threshold}}$ would increase in energy due to collisions with the thermal bulk. Imposing the requirement that the power-law start at Lorentz factors greater than the threshold Lorentz factor means that γ_1 will always be to the right of the thermal peak, thereby avoiding getting trapped at solutions with $\gamma_1 \rightarrow 1$.

Often, the cooling times are shorter than the energy loss times for all values of Lorentz factor. In this case, collisions are never important for the nonthermal electrons (Figure 4.2). However, it is still important

to maintain a thermal population of electrons to provide the bulk of the density in the plunging region. When no solution to Equation 4.15 exists, we set $\gamma_1 = \max(\gamma_{1,\text{threshold}}, \gamma_{\text{min}})$.

4.3.2.2 Determining A_{PL}

The power-law normalization A_{PL} is set by requiring continuity of the thermal and nonthermal electrons at γ_1 :

$$f_{\text{MJ}}(\gamma_1) = f_{\text{PL}}(\gamma_1). \quad (4.16)$$

4.3.2.3 Determining p_{eff}

The power-law index p_{eff} of the steady-state nonthermal distribution is calculated by assuming that none of the highest-energy particles are thermalized by Coulomb collisions; that is,

$$f_{\text{PL},0}(\gamma_2) = f_{\text{PL}}(\gamma_2). \quad (4.17)$$

We use an effective power-law index because high-energy particles collide more slowly than lower-energy particles, resulting in a power-law index that decreases over time [160]. Using the power-law index from PIC simulations would underestimate the total power law. Our formulation assumes that the initial power-law is continually replenished by in-situ particle acceleration, so the impact of cooling is primarily in the radiation rather than the dynamics. The difference between p_{eff} and p is largest close to the event horizon (Figure 4.3). Figure 4.4 shows how the final power-law distribution (blue solid line) differs from the initial power-law distribution (purple dashed line).

Combining Eqns. 4.12 and 4.17 yields an expression for p_{eff} :

$$p_{\text{eff}} = \frac{\log [f_{\text{PL},0}(\gamma_2)/f_{\text{MJ}}(\gamma_1)]}{\log [\gamma_1/\gamma_2]}. \quad (4.18)$$

4.3.2.4 Determining T_e

The temperature T_e of the thermal electrons is set by requiring a steady state that balances energy lost to radiation with energy dissipated by magnetic torques. We obtain the volume heating rate Q_+ from the

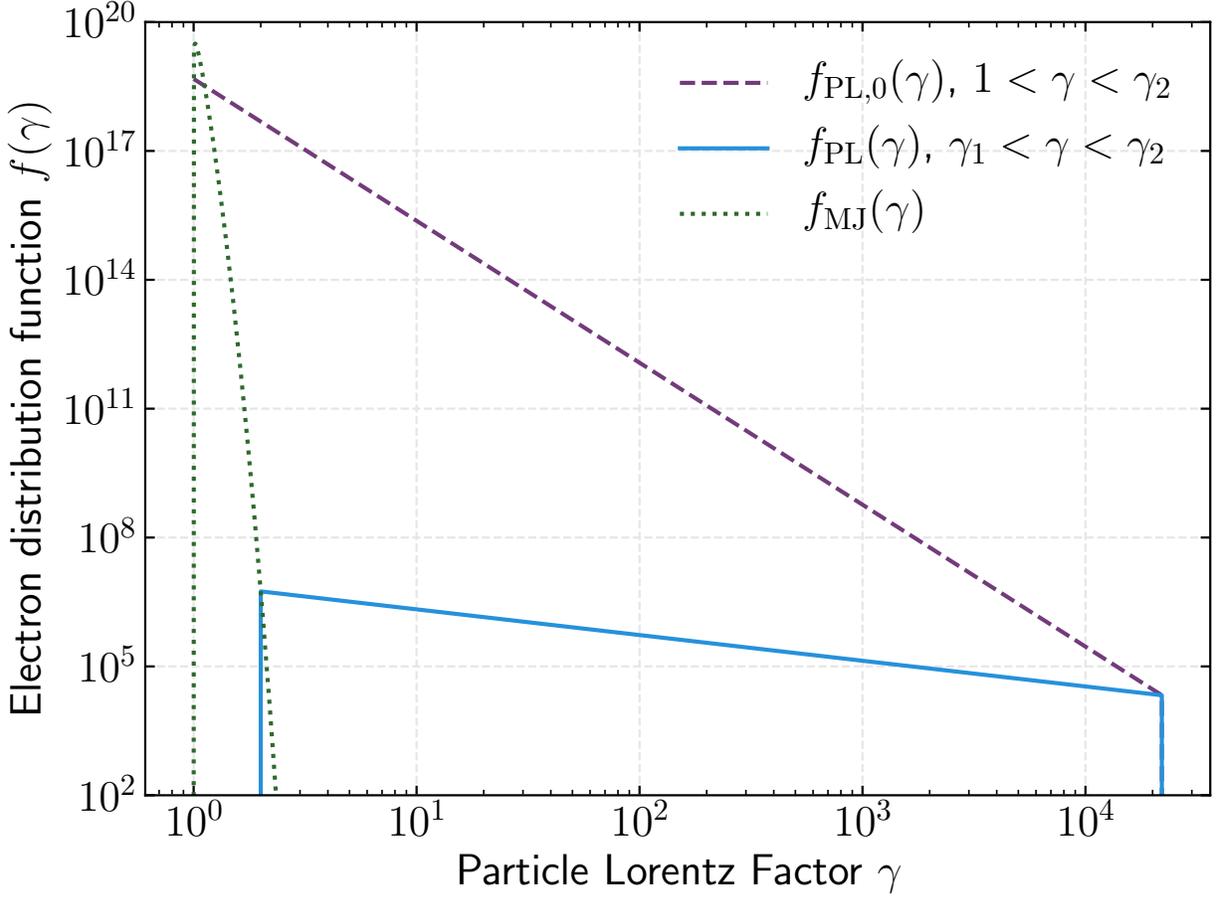


Figure 4.4: The effective power-law index p_{eff} is calculated from the initial power-law distribution $f_{\text{PL},0}$ by assuming that both power laws have the same value at the high-energy cut-off γ_2 (Equation 4.17). This sample electron distribution was calculated for fiducial parameters close to the event horizon at a radius $r = 1.31 r_g$, where $p = 3.3$ (purple dashed line) and $p_{\text{eff}} = 0.6$ (blue solid line). For radii closer to the ISCO, the difference between $f_{\text{PL},0}$ and f_{PL} is not so pronounced. The thermal distribution is also shown for reference (green dotted line).

accretion efficiency $\Delta\epsilon$ given by the Ref. [51] model, assuming that dissipation occurs evenly throughout the plunging region's volume:

$$Q_+ = \delta_e \frac{\Delta\epsilon L_{\text{Edd}}}{V_{\text{PR}}}, \quad (4.19)$$

where $V_{\text{PR}} = 4\pi(r_{\text{ISCO}}^3 - r_{\text{EH}}^3)/3$ is the (Newtonian) volume of the plunging region. The fraction of dissipated energy δ_e that goes into electrons is set at a constant value of 0.5, as motivated by PIC simulations of the energy partition in an electron-ion plasma with magnetizations close to 1 [68, 89]. The remainder $1 - \delta_e$ fraction of dissipated energy heats the ions, which do not radiate.

The total cooling rate Q_- sums the bremsstrahlung cooling rate $Q_{-, \text{MJ}}$ from the thermal electrons with the inverse Compton or synchrotron cooling rate from the nonthermal electrons. To account for additional scattering, the nonthermal cooling rate $Q_{-, \text{PL}}$ is set to the nonthermal emitted power $\mathcal{P}_{-, \text{PL}}$ decreased by a factor of $\tau_e = n_0 \sigma_T H = n_0 \sigma_T (H/r)r$, such that $Q_{-, \text{PL}} = \mathcal{P}_{-, \text{PL}}/\tau_e$. This additional factor estimates how many of the synchrotron/IC photons would escape the plunging region without additional scattering, which is important since $\tau_e \gtrsim 1$. This choice is motivated by Monte Carlo simulations of an isotropic sphere of gas with photons packets initialized evenly throughout the sphere. These simulations show that about $1/\tau_e$ of the electrons escape with no scattering. The remaining $1 - 1/\tau_e$ fraction of photons will Compton scatter off of the thermal/nonthermal electrons, heat the gas, and re-energize nonthermal electrons. The equation for steady-state thus reads:

$$Q_+ = Q_- = Q_{-, \text{MJ}} + Q_{-, \text{PL}}. \quad (4.20)$$

The nonthermal emissivity $\mathcal{P}_{-, \text{PL}}$ is calculated by integrating the Larmor formula for power emitted per electron over frequency ω , the power-law electron distribution $f_{\text{PL}}(\gamma)$, and solid angle. In integrating over solid angle Ω , we assume an isotropic distribution of pitch-angles α . The resulting power lost per unit volume is given by:

$$\mathcal{P}_{-, \text{PL}}^{\text{sync}} = \int d\Omega d\omega d\gamma f_{\text{PL}}(\gamma) \frac{2e^4 B_r^2 \gamma^2 \beta^2 \sin^2 \alpha}{3m_e^2 c^3} \quad (4.21)$$

$$= \frac{16\pi e^4 B_r^2}{9m_e^2 c^3} \int_{\gamma_1}^{\gamma_2} (\gamma^2 - 1) A_{\text{PL}} n_0 \gamma^{-p_{\text{eff}}} d\gamma. \quad (4.22)$$

If inverse Compton dominates, the emissivity is set to $\mathcal{P}_{-, \text{PL}}^{\text{IC}} = (U_{\text{ph}}/U_B) \mathcal{P}_{-, \text{PL}}^{\text{sync}}$.

We describe the cooling due to thermal particles in the plunging region as saturated Comptonized bremsstrahlung. Because the optical depth is of order 1 and the Compton y parameter is much greater than 1, the bremsstrahlung is amplified by repeated inverse Compton scatterings. The resulting emissivity is

$$Q_{-,MJ} = A(n_0, T_e) \epsilon_{ff}(n_0, T) \quad (4.23)$$

where $A(n_0, T_e) = 0.74 [\ln(2.25/x_{coh})]^2$ is an approximate amplification factor due to repeated inverse Compton scattering, $x_{coh} = 2.4 \times 10^{17} (m_e n_0)^{1/2} T_e^{-9/4}$ and $\epsilon_{ff}(n_0, T) = 1.68 \times 10^{-27} T_e^{1/2} n_0^2$ is the standard bremsstrahlung emissivity [57].

Because the energy loss collision times (Equation 4.7) assume a nonrelativistic thermal bulk, we put an upper limit on T_e at the temperature T_{\max} where setting $v_e^{ee}(x_c^{ee}) = 0$ would require a particle velocity greater than the speed of light. This maximum temperature is $T_{\max} = 3 \times 10^9$ K.

Equations 4.15, 4.17, 4.16, and 4.20 are solved iteratively at each radius using Brent's method to obtain $\gamma_1(r)$, $p_{\text{eff}}(r)$, $A_{\text{PL}}(r)$, and $T_e(r)$.

4.4 Results

4.4.1 High-spin, Highly-Magnetized Case

To illustrate the model, we examine the fiducial case $a = 0.95$ and $\Delta\epsilon = 1.04$, which corresponds to $F_{\theta\phi} = 6.0$ and magnetic flux a factor of 2 below the saturation value for a magnetically-arrested disk [156]. The hierarchy of timescales for this fiducial model justifies the assumptions outlined in Section 4.2. For this set of parameters, γ_1 always hits the lower bound $\gamma_{1,\min} = 2.0$. The equilibrium electron temperature θ_e decreases from 0.14 (8.9×10^8 K) at the ISCO to 0.03 (1.9×10^8 K) at the event horizon (Figure 4.3). These temperatures are roughly consistent with the inner disk temperatures in two-temperature GRMHD simulations [80]. The temperature decreases towards the event horizon because of the increase in γ_2 , decreasing the power law's overall normalization and shifting the temperature lower because of the continuity requirement (Equation 4.16). The decoupling of electrons and protons happens very close to the ISCO, at $1.82 r_g$, meaning that almost the entire plunging region has decoupled. The cooling rate is dominated by nonthermal synchrotron emission, with thermal cooling representing $\lesssim 1\%$ of the cooling rate at the ISCO

and 0.1% at the event horizon (Figure 4.3). This increase in the fraction of cooling from nonthermal electrons is again driven by the increase in γ_2 , which is set by the background magnetic field and gas density.

The total spectrum, shown in Figure 4.5, consists of three parts: the plunging region's nonthermal electrons (blue solid line), the plunging region's thermal electrons (green dotted line), and the thin disk's spectrum (orange dashed line). The spectrum is calculated with redshifts obtained by ray tracing the plunging region-disk system with the code `geokerr` to a camera inclined at $i = 60^\circ$ to the black hole and disk angular momentum vector [162]. The gas is assumed to have a four-velocity given by the Ref. [51] model within the plunging region and a purely azimuthal, Keplerian velocity in the disk body. The power law intensities $I_{\nu, \text{PL}, ij} = P_{\nu, \text{PL}, ij} H / (4\pi)$ at a camera pixel with position i and j in the x - and y -directions are calculated by assuming a hybrid electron distribution at the radius r_{ij} where the ray-traced photon hits the plunging region midplane. The power-law distribution is $P_{\nu, \text{PL}, ij} = A_{ij} \nu^{-s_{ij}}$ between the characteristic synchrotron frequencies $\nu_{1, ij}$ and $\nu_{2, ij}$ for electrons with Lorentz factors $\gamma_{1, ij}$ and $\gamma_{2, ij}$. That is, $\nu_1 = \nu_0 \gamma_{1, ij}^2$ and $\nu_2 = \nu_0 \gamma_{2, ij}^2$, where $\nu_0 = 3e|B_r| / (4\pi m_e c)$. The power-law slope is $s_{ij} = (p_{\text{eff}, ij} - 1) / 2$. The constant A_{ij} is set by requiring that $\int P_{\nu, \text{PL}, ij} d\nu = Q_{-, \text{PL}, ij}$. The emission from the thermal plunging region electrons is in the Comptonized Wien regime, such that $I_{\nu, \text{MJ}, ij} = I_{\nu}^W(T_{ij}) e^{-\alpha_{ij}}$, where T_{ij} is the thermal electron temperature at r_{ij} , $I_{\nu}^W(T_{ij}) = 2h\nu^3 / c^2 e^{-kT_{ij}/h\nu}$ is the Wien intensity, and $e^{-\alpha_{ij}}$ is the Compton saturation factor. The Compton saturation factor is determined by setting the frequency-integrated intensity equal to the isotropic thermal Comptonized bremsstrahlung emissivity: $I^W(T_{ij}) e^{-\alpha_{ij}} = Q_{-, \text{MJ}} H / 4\pi$. Luminosities are obtained by summing over 64×64 camera pixels: $L_{\nu} = 4\pi \Delta\alpha \Delta\beta \sum_{ij} I_{\nu, ij}$, where $\Delta\alpha = \Delta\beta = 0.125 r_g$ are the pixel widths. The disk spectrum is calculated by assuming a blackbody with a radially-dependent temperature extending from the ISCO to 10^{10} cm [46]:

$$T(r) = \left(\frac{3GM\dot{M}}{8\pi\sigma r^3} \frac{\mathcal{Q}}{\mathcal{B}\mathcal{C}^{1/2}} \right)^{1/4} \quad (4.24)$$

where \mathcal{B} , \mathcal{C} , and \mathcal{Q} are relativistic corrections [49, 163].

Several features of the spectrum in Figure 4.5 stand out. First, emission from the nonthermal plunging region electrons dominates the emission above 10 keV and produces an observable high-energy power-law tail. The power law has a relatively flat slope past 200 keV, a feature that is difficult to produce through

thermal Comptonization [87]. This power law cuts off at a little less than 100 MeV, though above 1 MeV pair production could become important (gray region). The total luminosity from the power law across all frequencies is 6.1% of the disk luminosity. This fraction decreases to 1.7% if non-X-ray frequencies below 1keV are excluded. Excluding luminosity from the pair-emitting frequency range above 1 MeV further decreases the fraction to 1.3%. The nonthermal power-law electrons dominate emission from the plunging region, with a luminosity 63, 15, or 12 times that of the thermal plunging region electrons for all frequencies, frequencies above 1 keV, and frequencies between 1keV and 1 MeV, respectively.

Most of the observed luminosity comes from radii close to the ISCO, as shown in Figure 4.6. This figure calculates the amount of observed luminosity that originates from between the event horizon and a radius r . We define the half-light radius $r_{1/2}$ as the location within which 50% of the light from the plunging region has been emitted; analogously, $r_{1/10}$ is where 10% of the light has been emitted. For this fiducial case with the ISCO at $r = 1.94r_g$, 50% of the light comes from $r > 1.77r_g$ and 90% comes from $r > 1.66r_g$.

The emission depends on the spin and the inclination of the disk-plunging region system. The inclination of the disk with respect to an observer changes the effective observed area, an effect that goes as $\cos i$, where i is the angle between the black hole spin axis and the line-of-sight. Relativistic beaming effects have the opposite trend with inclination angle; radiation from relativistically-moving gas is beamed into the plane of motion, leading to a higher luminosity for larger inclination angles. In the plunging region, beaming effects dominate, resulting in a factor of 100 increase in power-law luminosity from $i = 0^\circ$ ($L_{\text{PL}} = 5 \times 10^{35}$ erg/s) to $i = 90^\circ$ ($L_{\text{PL}} = 6 \times 10^{37}$ erg/s) for $a = 0.95$. Power-law emission for $a = 0.5$ is roughly constant over inclination angle, with $L_{\text{PL}} \sim 2 - 3 \times 10^{36}$ erg/s. Beaming effects no longer dominate for the smaller spin because the fluid Lorentz factor is smaller due to a larger r_{ISCO} . The blackbody disk's peak temperature depends on the location of the ISCO, which is set by the black hole spin and decreases by a factor of about 3 from a non-spinning to a maximally-spinning black hole [164, 165]. The disk emission is non-monotonic as a function of inclination angle, peaking around 60° , where area projection effects start to dominate over Doppler effects. The peak disk emission is a factor of two larger than at $i = 0^\circ$ for $a = 0.95$, and a factor of 1.7 larger for $a = 0.5$. As a result, $L_{\text{PL}}/L_{\text{disk}}(1\text{keV} < \nu < 1\text{MeV})$ for $a = 0.95$ increases from 0.002 at $i = 0^\circ$ to 0.17 at $i = 85^\circ$, whereas for $a = 0.5$ it increases from 0.055 at $i = 0^\circ$ to 0.09 at $i = 85^\circ$, with a minimum

of 0.042 at $i = 45^\circ$.

4.4.2 Parameter Space

Figure 4.7 shows the power law to disk luminosity fraction as a function of plunging region dissipation and black hole spin. For $i = 60^\circ$, the plunging region can emit up to 1% of the accretion disk luminosity for a high-spin black hole, and even higher for lower spin black holes. At constant spin, disks with less magnetic torque at the ISCO and hence a lower accretion efficiency $\Delta\epsilon$ have less nonthermal luminosity because weaker magnetic fields lead to less efficient cooling of high-energy particles, whereas the thin disk emission is constant with $\Delta\epsilon$. At fixed $\Delta\epsilon$, lower black hole spins lead to a more visible power law because the plunging region extends to larger radii, increasing the magnitude of the nonthermal plunging region luminosity and decreasing the disk luminosity. For $\Delta\epsilon \approx 0.4$, the power law's X-ray luminosity increases by a factor of 400, from 10^{33} erg/s for $a = 0.95$ to 4×10^{35} erg/s for $a = 0.70$. This increase is due to a combination of a larger area radiating 90% of the emission and redshifts closer to unity for smaller spins. The thin disk's X-ray luminosity decreases for decreasing spin by a factor of ten, from 10^{38} to 10^{37} erg/s. These combined effects explain why $L_{\text{PL}}/L_{\text{disk}}$ is a factor of 4000 different for the same magnetization.

Our model demonstrates two regimes for the plunging region X-ray emission: one where nonthermal electrons dominate and one where thermal electrons dominate. For fixed spin, weaker magnetic fields lead to less cooling by nonthermal electrons and thus require more thermal cooling and higher thermal temperatures to balance the total heating rate. Higher thermal temperatures and a less prominent power-law lead to a significant decrease in $L_{\text{PL}}/L_{\text{MJ}}$ with decreasing accretion efficiency. Nonthermal electrons dominate only at the highest disk magnetizations because particle acceleration requires strong magnetic fields. For less strongly magnetized disks, thermal electrons dominate the plunging region emission; in this regime, we might expect the observed power law to come from thermal electrons with a rapidly-increasing temperature rather than nonthermal electrons [157]. Figure 4.7 marks the regime where nonthermal electrons dominate with black circles around the colored markers.

To facilitate comparison with observations, we show our model's predictions for the power-law fraction

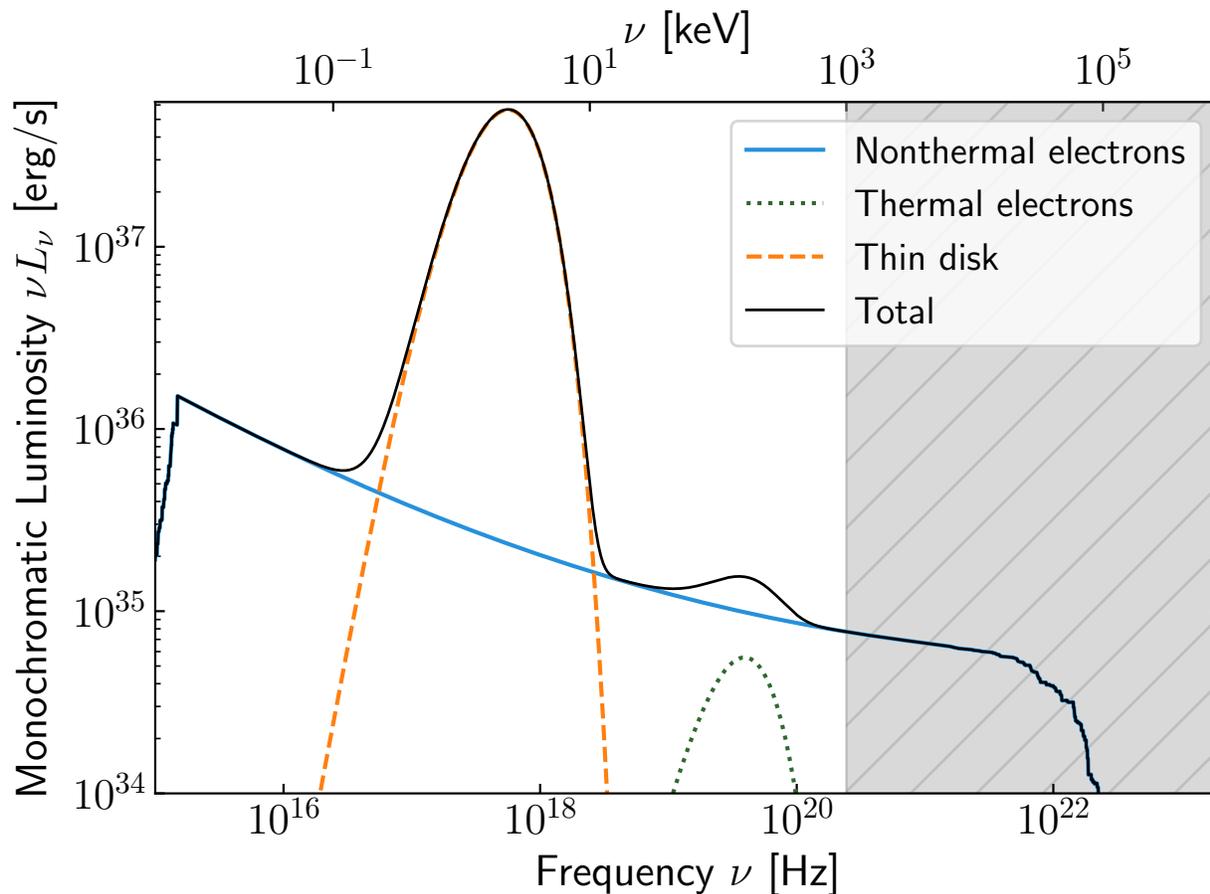


Figure 4.5: Spectrum for the fiducial model parameters $a = 0.95$ and $F_{\theta\phi} = 6.0$ at an inclination of 60° . The thin disk emission (dashed orange line) peaks around a few keV. The nonthermal electrons' radiation (solid blue line) extends from 10 eV to 0.1 GeV, although emission beyond 1 MeV (gray region) may be impacted by pair creation. The thermal electrons in the plunging region create a small excess at around 10 keV (dotted green line).

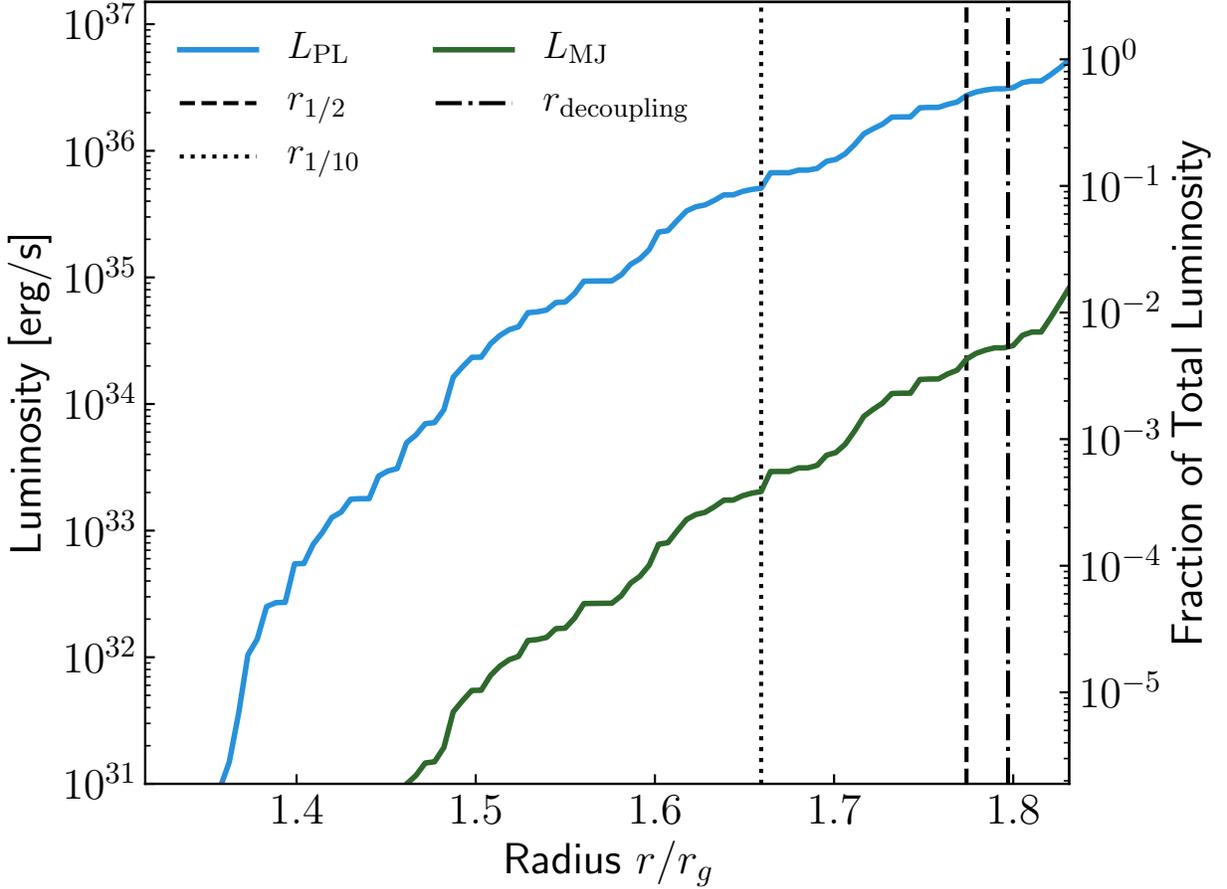


Figure 4.6: Portion of the fiducial model’s observed frequency-integrated luminosity originating between the event horizon and a radius r for an inclination angle of 60° . Emission from the plunging region’s nonthermal electrons (blue) dominates over the plunging region’s thermal electrons (green). 10% of the luminosity comes from $r < r_{1/10} = 1.66r_g$ (dotted line), while 50% comes from $r < r_{1/2} = 1.77r_g$ (dashed line). Vertical dash-dot line shows the decoupling radius.

(PLF) as a function of inclination angle and spin in Figure 4.8. The PLF is defined as

$$\text{PLF} = \frac{\mathcal{L}_{\text{PL}}}{\mathcal{L}_{\text{PL}} + \mathcal{L}_{\text{disk}}} \quad (4.25)$$

where $\mathcal{L}_{\text{PL}} = L_{\text{PL}}(1 \text{ keV} < \nu < 100 \text{ keV})$ is the power-law luminosity in the X-rays and $\mathcal{L}_{\text{disk}} = L_{\text{disk}}(0.001 \text{ keV} < \nu < 100 \text{ keV})$ is the disk luminosity over a wide range of frequencies, to avoid cutting off low-spin blackbody emission that peaks at $\sim 1 \text{ keV}$ [166, 167]. For the high-spin case, the PLF is almost the same as $L_{\text{PL}}/L_{\text{disk}}(1 \text{ keV} < \nu < 1 \text{ MeV})$ because $L_{\text{PL}} \ll L_{\text{disk}}$. For lower spins, the PLF is lower than $L_{\text{PL}}/L_{\text{disk}}(1 \text{ keV} < \nu < 1 \text{ MeV})$ by a factor of 2 - 3 because of the change in denominator. The parameters in Figure 4.8 are the most strongly magnetized disks where the nonthermal emission dominates thermal emission from the plunging region; for the least magnetized cases, the PLF is less than 10^{-3} (not shown) and thermal plunging region electrons would significantly change the shape of the emission above 10 keV. Overall, the PLF calculated from the model lies between 10^{-3} and 0.3, values consistent with observations of the soft state [166].

4.5 Discussion

4.5.1 Application to Astrophysical Systems

Our model broadly agrees with observations of X-ray binaries in the soft state. The soft state usually exhibits a photon index $\Gamma \gtrsim 2$ [48], corresponding to model fits with an electron power-law index $p \gtrsim 3$. The model's most strongly magnetized parameters give an electron power-law index of $3 - 4$ (Figure 4.3), which increases for decreasing magnetization. The soft state's gamma-ray tail has luminosities on the order of $10^{35} - 10^{37} \text{ erg/s}$ above 50 keV [168]. Our model finds power-law luminosities above 50 keV between $10^{34} - 10^{37} \text{ erg/s}$, within the same order of magnitude. Calculations from observations find that the PLF varies between 10^{-3} and 0.2 in the soft state of XRBs [166], which agrees well with the values found from our model (Figure 4.8). As yet, trends of the PLF with inclination angle and spin cannot be found from observations [167]. Testing our predictions concretely against observations will require less uncertainty in the PLF.

The power-law emission in our model comes from synchrotron emission, whereas typical models for

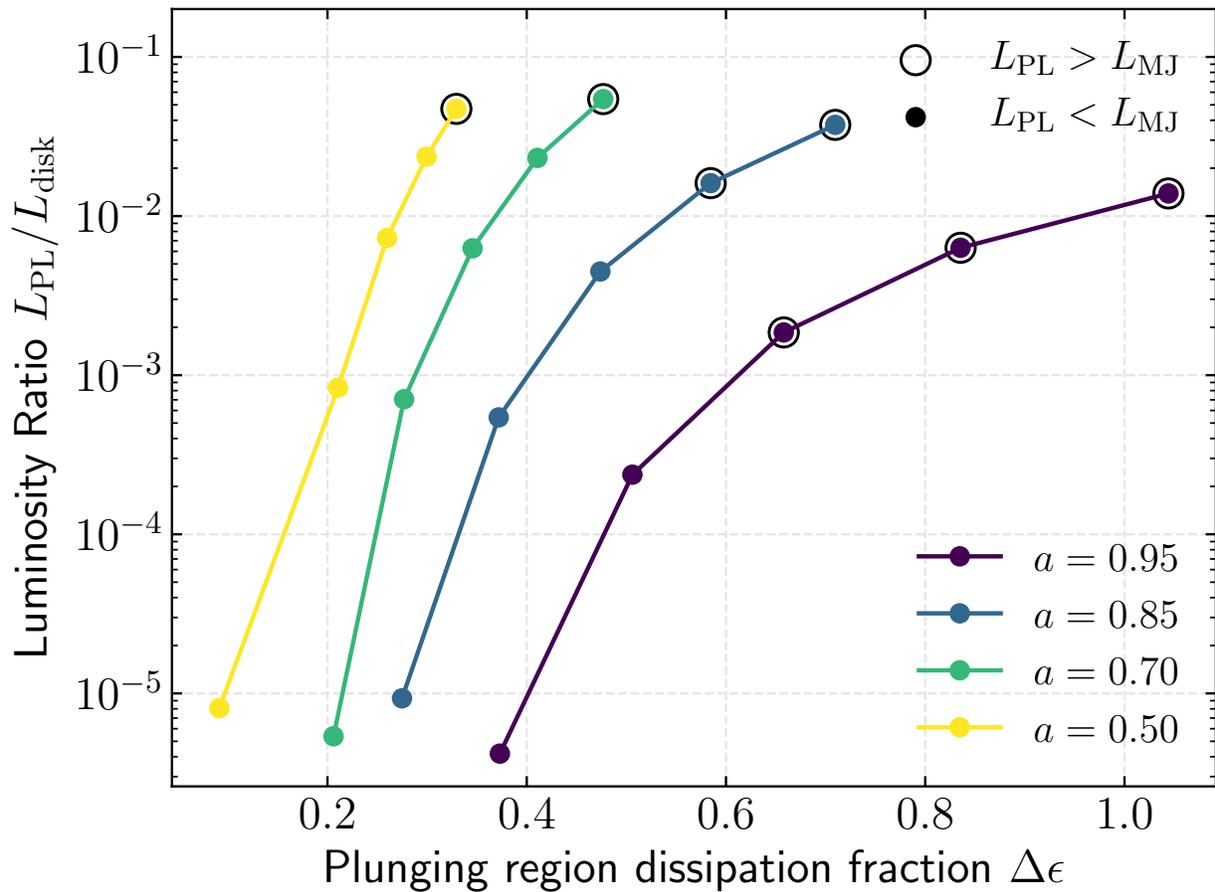


Figure 4.7: Frequency-integrated luminosity in X-rays ($1 \text{ keV} < \nu < 1 \text{ MeV}$) in predicted power law compared to the thermal blackbody disk for different model parameters, assuming an inclination angle $i = 60^\circ$. Models where nonthermal electron emission dominates over thermal electron emission, satisfying $L_{\text{PL}} > L_{\text{MJ}}$, are marked with large black circles.

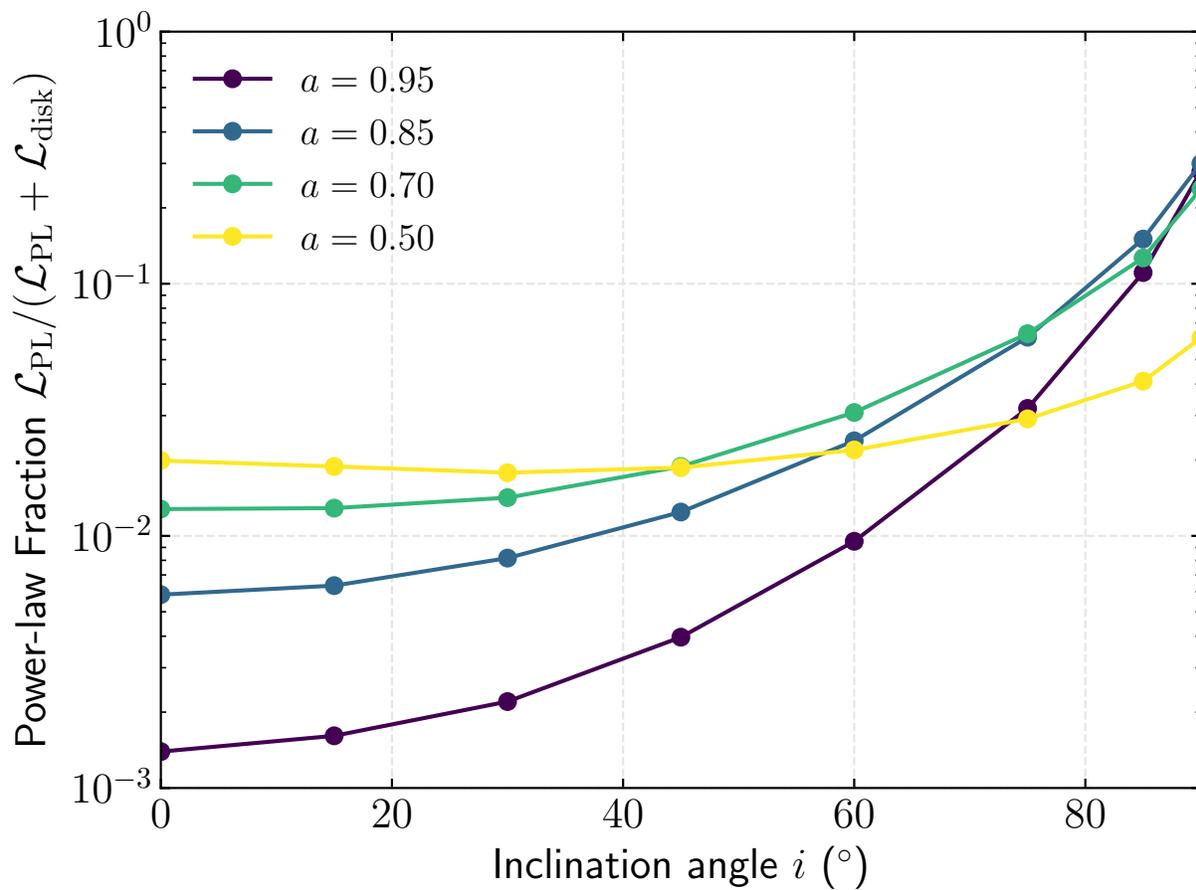


Figure 4.8: Power-law fraction (Equation 4.25) as a function of inclination angle. Calculated from the most strongly magnetized plunging region parameters such that nonthermal plunging region emission dominates thermal plunging region emission.

the power law assume inverse Compton scattering of blackbody disk photons off nonthermal particles [149, 152]. Our model therefore predicts more strongly polarized emission for the soft state's emission above 10 keV than other models where inverse Compton scattering dominates. Current polarization measurements set an upper limit of 70% for Cyg X-1's 0.4 - 2 MeV emission in the soft state [169], which does not constrain the radiation mechanism.

The hard state of XRBs such as Cyg X-1 shows an excess above 400 keV that fits well to a hybrid thermal/nonthermal model with a power-law electron distribution index between 3.5 and 5 [170]. The nonthermal electrons in that model could also be accelerated by turbulence or magnetic reconnection. If the hierarchy of timescales satisfies Equation 4.10, the multizone equilibrium model presented in Section 4.3 could be adapted to a geometry relevant to the hard state.

4.5.2 Model Limitations

Many of our model assumptions were motivated by GRMHD simulations. Our finding that ions and electrons decouple in the plunging region is consistent with estimates of the thermalization time from single-temperature, ideal GRMHD simulations [157]. In fact, the decoupling region could extend even outside the ISCO. The Ref. [51] model artificially sets the radial velocity to zero at the ISCO, forcing a thin disk at $r > r_{\text{ISCO}}$ whereas the thin disk could in principle thicken at larger radii depending on the dominance of Coulomb collisions. For simplicity, in our model, the ion temperature and thus the ratio H/R was discontinuous at the ISCO. This discontinuity is not present in full 3D GRMHD simulations; instead, a transition region forms between a thin disk and the hot flow [79, 80]. Such a transition region would likely result in a less distinct bump from the thermal electrons in the plunging region. GRMHD simulations also motivate magnetic reconnection as an important dissipation mechanism, demonstrating hallmarks of reconnection such as current sheets and plasmoids throughout a magnetically-dominated disk [171, 172]. Our prescription for γ_2 and p does not include the impact of particle anisotropy or magnetic guide field on the electron distribution [90, 173], which could alter the observed spectrum. Including a guide field tends to steepen p and decrease γ_2 for a pair plasma [85, 90]. Pitch-angle anisotropy in a pair plasma can also lead to temperature anisotropy for the thermal bulk [116, 173], which could affect radiation from the thermal

plunging region electrons. Unlike GRMHD simulations, our model does not include vertical structure and we assumed a constant $H/R = 1$ and a uniform distribution of the dissipation throughout the plunging region volume. Vertical structure would presumably lead to a dense equatorial region with more optically thin upper parts of the plunging region. We also assume uniform dissipation such that $Q_+ \propto \rho$, as commonly done in disk atmosphere models [174].

For simplicity, we did not include full radiative transport in our model. Instead, we assumed that the thermal plunging region electrons cooled solely through thermal bremsstrahlung, neglecting thermal synchrotron radiation because the average Lorentz factor of the thermal particles was less than 2 (i.e. temperatures less than $\sim 10^9$ K). We also neglect synchrotron self-absorption, since this effect is only important for frequencies below $\sim 10^{15}$ Hz [175], outside the range we consider. We neglect cyclotron radiation for the same reason. We restrict the observed radiation from the nonthermal electrons to be either from synchrotron radiation or from inverse Comptonization of soft disk photons, i.e. we neglect synchrotron self-Compton emission (SSC). Neglecting SSC does not result in large errors for small optical depth and spectral index > 1 , where synchrotron emission dominates over SSC [176]. We also do not consider synchrotron and inverse Compton processes simultaneously; instead, we assume one or the other dominates. We also assume a simple thermal + power-law electron distribution function, whereas Fokker-Planck models evolving a power-law electron distribution thermalizing under electron-electron collisions show a more complicated structure [160]. Although pair production could become important in this regime, it remains unclear how pair-regulated reconnection would change when synchrotron losses are included [177]. We note that Compton reflection off the cold disk could increase the luminosity above 10 keV by order unity for sandwich-type models [150]. Including Compton reflection of the plunging region radiation off the cold disk could also increase the high-energy luminosity and lead to higher power-law fractions. More detailed modelling of reflection features such as the Fe $K\alpha$ emission line is beyond the scope of this work.

4.6 Conclusions

We have presented a semi-analytic model for the soft state of XRBs where the high-energy power-law tail is produced by nonthermal electrons in the plunging region. We demonstrated the feasibility of

having nonthermal electrons in the plunging region by examining the hierarchy between electron-proton thermalization time, electron-electron collision time, electron cooling time, electron acceleration time, and the infall time (Section 4.2, Figure 4.2). Using an analytic dynamical background and the results of particle-in-cell simulations of magnetic reconnection, we constructed a steady-state model for the electron distribution function at each radius in the plunging region (Section 4.3). The nonthermal electrons in this model produce an observable power law from 10 – 1000 keV with a photon spectral index $\Gamma \gtrsim 2$ (Figure 4.5). By exploring the model parameter space, we show that plunging region emission from nonthermal electrons dominates over thermal electrons for strongly-magnetized models and vice-versa for less strongly-magnetized models, suggesting a Ref. [157]-type model. We found fractions of power law to total X-ray luminosity (power-law fraction, PLF) consistent with observations for all values of spin (Figure 4.7). We predict an increase in the observed PLF with inclination angle (Figure 4.8). Although the PLF in our model is consistent with observational values, testing the trend with inclination angle and spin will require more observations.

Future work using numerical simulations could explore how and where the decoupling happens, as well as its possible role in spectral state transitions and jet/wind launching.

Acknowledgements

The authors thank Mitch Begelman, Dmitri Uzdensky, and the referee, Luca Comisso, for helpful comments. AMH acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1650115. This work was supported in part by NASA Astrophysics Theory Program grants NNX16AI40G, NNX17AK55G, and 80NSSC20K0527 and by an Alfred P. Sloan Research Fellowship (JD).

Data Availability

The simulation data underlying this article will be shared on reasonable request to the corresponding author.

Chapter 5

GRMHD Simulations of Accretion Disk Truncation due to Coulomb Collisions

In this chapter, I build on Ch. 4 to study the impact of Coulomb collisions on the structure of the disk-corona system in a global sense. The geometry of the disk-corona system remains largely unknown, although some observations have suggested a relatively compact corona. Fully 3D GRMHD simulations often assume the accretion disk plasma has a single temperature. The single-temperature assumption does not hold for the thick disks used by the disk truncation model to describe the inner parts of the accretion flow during the hard state of black hole binaries. Having previously analytically calculated that electrons and protons should decouple within the plunging region of a thin disk, I now investigate how Coulomb collisions can shape the disk both outside and inside the ISCO. Previous work has suggested that Coulomb collisions due to a changing accretion rate could lead to the collapse of a thick disk and thus a state transition. However, modeling radiation in the intermediate accretion regime is quite challenging numerically. In this work, I model radiation as a cooling function that only affects electrons. I aim to describe the transition between a thin and thick disk (if it exists) as a function of magnetic field and accretion rate. This work is ongoing.

5.1 Introduction

In accretion disk theory, optical depth characterizes the zeroth-order type of accretion flow. The standard thin disk has $\tau \gg 1$ [9], whereas the well-known advection-dominated accretion flow (ADAF) is optically-thin [56]. Transitions between these two accretion modes could occur in X-ray binaries, whose spectral states are consistent with a puffy, optically thin flow in the hard state and a thin, optically thick disk in the soft state [48]. Recent numerical simulations suggest that the hard-to-soft spectral transition (and thus

thick-to-thin disk transition) could be triggered due to the efficiency of Coulomb collisions between electrons and protons at intermediate mass accretion rates, collapsing the thick disk [58]. However, the numerical cost of the Monte Carlo radiation prevents the disk from evolving long enough to actually collapse. As tends to be the case, the numerically challenging regime is also one of the most interesting regimes.

Treating radiation in accretion disks across optical depth regimes poses a challenge to existing numerical methods in general relativistic magnetohydrodynamic (GRMHD) codes. In the optically thin regime where the optical depth τ is much less than 1, Monte Carlo methods can capture the relevant physics [99, 100, 104]. On the other hand, in the optically thick regime of $\tau \gg 1$ where scattering can be considered local, treating radiation as a fluid using an M1-closure suffices [101]. However, both methods have disadvantages when moving to the opposite regime. The M1-closure suffers from problems when scattering becomes highly nonlocal, whereas Monte Carlo methods become prohibitively expensive as the optical depth increases and the number of scatterings of superphotons increases.

The transition from optically-thin to optically-thick flow has been modelled numerically a handful of times. Introducing a bistable cooling function leads to a thick-to-thin transition [79], but of course depends explicitly on the location of the temperature switch, which is put in by hand. More recently, a single simulation using the M1-closure method showed the self-consistent development of an inner hot flow and an outer thin disk around $20 GM/c^2$ for a two-temperature plasma [80]. The truncation in this case was due to magnetic torques in a strongly magnetized disk.

Our goal is to develop a method that inexpensively captures the impact of radiation while not assuming strong Coulomb coupling, thereby allowing for a thick-to-thin disk transition. To that end, we propose an *electron-only* cooling function. Cooling only the electrons emulates the fact that protons do not cool radiatively, and will only cool if Coulomb collisions with electrons are efficient. If we assume that synchrotron radiation or inverse Compton scattering is the dominant cooling mechanism in the inner disk [5], then relativistic electrons should rapidly lose energy until they reach Lorentz factors $\gamma \sim 1$, i.e. a temperature of $\sim 10^9$ K. We will therefore begin by setting a target temperature $T_{e,\text{target}} = 10^9$ K, although in principle this number can vary — and will, once synchrotron cooling no longer dominates and the plasma is no longer optically thin.

5.2 Methods

For demonstration purposes, consider the public GRMHD code `ebhlight` [99, 100] that solves the general relativistic ideal magnetohydrodynamic equations using the HARM scheme [98, 178]; see Ch. 2.5.1. Conservative codes such as `ebhlight` conserve their primitive variables, in particular the total energy and total entropy. The total energy u_g includes the proton and electron energies u_p and u_e respectively, which sum together linearly:

$$u_g = u_p + u_e. \quad (5.1)$$

Protons and electrons have different adiabatic indices: electrons are relativistic so $\gamma_e = 4/3$, whereas the protons are nonrelativistic and have $\gamma_p = 5/3$. Because of this difference in adiabatic indices, the electrons and protons in a given volume heat (or cool) by different amounts when the volume compresses (or expands). This difference in heating follows from the relationship between internal energy density and temperature:

$$P_k = (\gamma_k - 1)u_k, \quad (5.2)$$

where the subscript k labels the species (total gas, electron, or proton) and P is the pressure. Therefore, from Eq. 5.1, the total gas temperature is not simply a sum of electron and proton temperature. Instead,

$$T_g = \frac{\gamma_g - 1}{\gamma_e - 1}T_e + \frac{\gamma_g - 1}{\gamma_p - 1}T_p. \quad (5.3)$$

Total entropy and electron entropy are also conserved. Here, we will refer to the quantity

$$\kappa_k = P_k \rho^{-\gamma_k} = (\gamma_k - 1)u_k \rho^{-\gamma_k} = e^{(\gamma_k - 1)s_k}, \quad (5.4)$$

where s_k is the entropy per particle. Using Eq. 5.1, 5.2, and 5.4, we can write the total gas entropy in terms of electron and proton entropy as:

$$\kappa_g = P_g \rho^{-\gamma_g} = (\gamma_g - 1)u_g \rho^{-\gamma_g} = (\gamma_g - 1) \left(\frac{P_e}{\gamma_e - 1} + \frac{P_p}{\gamma_p - 1} \right) \rho^{-\gamma_g} \quad (5.5)$$

$$= \frac{\gamma_g - 1}{\gamma_e - 1} \rho^{\gamma_e - \gamma_g} \kappa_e + \frac{\gamma_g - 1}{\gamma_p - 1} \rho^{\gamma_p - \gamma_g} \kappa_p. \quad (5.6)$$

5.2.1 Electron Fluid

Modelling electrons separately from protons is important for capturing radiative effects accurately. Previous implementations either assume that the electron temperature is the same as the gas temperature, which is completely invalid in the low density gas, or assume a fixed ratio T_p/T_e that depends on variables such as radial location or magnetic field [179, 180]. However, these models do not actually keep track of the electron heating or cooling. In order to use prescriptions from PIC simulations (see Ch. 2.5.2) that determine what fraction of viscous heating goes into protons vs. electrons from magnetic reconnection or turbulence [67, 85, 86], simulations must actually keep track of electrons separately from protons.

An electron fluid, evolved separately from the background fluid [103], heats via Coulomb collisions and grid-scale dissipation. Codes like `ebhlight` assume that the electrons do not affect the fluid dynamics, only the thermodynamics. The electrons heat via grid-scale viscous dissipation. The viscous dissipation is calculated from the difference of the conserved total entropy κ_g and the entropy calculated from the conserved gas internal energy u_g , i.e.

$$\hat{\kappa}_g = (\gamma_g - 1)u_g\rho^{-\gamma_g}, \quad (5.7)$$

where $\hat{\kappa}_g$ includes numerical viscosity and resistivity. Because the stress tensor is conserved to machine precision, the numerical viscosity/resistivity appears in the code as entropy, $\hat{\kappa}_g - \kappa_g$. This entropy comes from the truncation errors in the code and as such only obeys the second law of thermodynamics over long time- and volume-averages. Because total energy u_g is conserved, entropy in this implementation can only occur when one form of energy converts to another, i.e. magnetic to kinetic energy.

With the introduction of an electron fluid, proton quantities are calculated as the difference between the total fluid and the electron fluid. Proton entropy, for example, is never explicitly evolved.

The code updates the electron entropy and total fluid entropy and outputs the viscous heating Q_{visc} for both electrons and protons. The total heating rate Q_{visc} is calculated from the difference between the entropy-conserving and the energy-conserving variables as described above. A fraction f_e of the total heating rate goes into electrons, while the remainder $(1 - f_e)Q_{\text{visc}}$ heats the protons. The fraction f_e is a source of much uncertainty. Many early models use an analytic prescription based off Landau damping of the magnetic

cascade [66]. Lately, PIC simulations have probed how f_e changes with magnetization, guide field, and temperature among other variables in both magnetic reconnection set-ups and magnetic turbulence [67, 68, 86, 181]. For this work, we will use the prescription from magnetic reconnection [68].

5.2.2 Electron-only Cooling Function

The cooling function described below affects only this electron fluid, in contrast to single-fluid GRMHD cooling functions that affect the entire gas, implicitly assuming strong Coulomb coupling.

The implementation of the electron cooling function follows Ref. [98]. We assume that any radiative cooling is optically thin, leaving the system without interacting with the gas. Assuming isotropic radiation in the fluid's rest frame, the radiation removes energy from the local energy conservation equation via:

$$G_\mu = -\mathcal{L}u_\mu, \quad (5.8)$$

where $G_\mu = \nabla_\nu T_\nu^\mu$, u_μ is the fluid four-velocity. The quantity \mathcal{L} is the energy radiated per unit time in the fluid's frame. We take

$$\mathcal{L} = \frac{2su_e Y^q}{t_{\text{cool}}}, \quad (5.9)$$

where u_e is the electron internal energy density and $Y = T_e/T_{e,\text{target}} - 1$. We typically set $q = 1$ and $s = 1$; see Ch. 5.3 for a parameter scan of these values. The cooling time t_{cool} depends on the problem parameters. For a GRMHD simulation, we typically set $t_{\text{cool}} = \Omega^{-1}$, where $\Omega = (r^{3/2} + a)^{-1}$ is the Keplerian orbital frequency. Inside the ISCO, we assume a constant angular momentum. Associating t_{cool} with the dynamical time provides a short timescale, but likely severely overestimates the actual cooling time, which should be orders of magnitude shorter. To avoid cooling the electrons in the unphysical jet region, we do not apply the cooling function in regions where the magnetization $\sigma \equiv B^2/\rho > \sigma_{\text{cut}}$. Unless otherwise noted, we take $\sigma_{\text{cut}} = 1$.

Checks and benchmarking of this electron-only cooling function are ongoing and will be described in the following sections.

5.2.3 Injected Viscous Heating Test

Viscous heating is a major source of heating in the torus problem. As such, I want to test how viscous heating, Coulomb heating, and electron cooling all interact in a controlled set-up, i.e. in my gas box. Ideally, I would inject a certain amount of viscosity that I prescribe and control. To illustrate some subtleties of the test and to explain some of the behavior I've been seeing in the gas box, I will draw on the previous descriptions of how the electron fluid and electron heating are implemented.

Remember that total energy u_g is conserved during the viscous heating process. From the right side of the second equals sign in Eq. 5.4, the conservation of u_g and ρ while κ_g increases means that γ_g should change. Ref. [103] acknowledges that assuming γ_g is constant does not hold if $\gamma_e \neq \gamma_p$, but argues that the assumption is reasonable because the change will not be significant.

Because the total gas temperature is determined from u_g , that also means that the total gas temperature T_g does not change during the viscous heating step. As a result, we can write the proton temperature change as a function of the electron temperature change using Eq. 5.3:

$$\Delta T_p = -\frac{(\gamma_p - 1)}{(\gamma_e - 1)} \Delta T_e = -2\Delta T_e. \quad (5.10)$$

The protons actually decrease in temperature during the viscous heating process. In gas box tests where I fixed $\Delta\kappa_g$, I confirmed that the proton temperature decreased by twice the amount the electron temperature increased. The lack of a total gas temperature increase combined with the forced electron temperature increase means that viscous heating will be completely counteracted by Coulomb energy exchange when the electrons and protons start with the same temperature.

5.3 Low Accretion Rates

We first test out the electron-only cooling function at low accretion rates where Coulomb coupling is essentially negligible. In these tests, we can examine how electrons heat (due to adiabatic compression only, no Coulomb heating) and cool due to “radiation”, i.e. the cooling function. These tests are 2D Fishbone-Moncrief tori with an inner radius of $6r_g$ and a resolution of 160x128 cells in the r and θ directions, respectively.

5.3.1 Parameter Scans

As shown in Fig. 5.1a's blue line, the canonical simulation parameters lead to an electron temperature that did not follow the target temperature. The electron temperature sits at the target of 10^9 K for radii larger than $r \gtrsim 4r_g$. Within this radius, which is close to the ISCO, the electrons do not cool down to the target, instead exceeding the target by a factor 7. One of the main topics of research for the first six months was diagnosing why the electrons did not effectively cool down to the target.

Fig. 5.1 shows how the electron thermodynamics depend on a variety of parameters. One of our best tools for controlling the electron thermodynamics is adjusting the cooling time. As shown in Fig. 5.1a, lowering the cooling time to e.g. $t_{\text{cool}} = 1/(2\Omega)$ does lower the temperature inside the ISCO, but does not completely eliminate the problem of the electron temperature being above the target. Even extremely short cooling times such as $t_{\text{cool}} = 1/(10\Omega)$ (not shown) did not fully reach the target temperature. These lower cooling times have an additional problem: they experience much more supercooling.

Another potential issue is the electron adiabatic index. Although the electrons are initially relativistically-hot with an adiabatic index of $\gamma_e = 4/3$, when they cool down to $T_{e,\text{target}} = 10^9$ K, they are no longer relativistically-hot and should have a nonrelativistic adiabatic index $\gamma_e = 5/3$. In the gas evolution, $P_e V_e^\gamma$ is constant and $P_e V \propto n T_e$ (ideal gas law). That means $T_e V^{\gamma_e - 1}$ is constant. So, for the same volume change V_1 to V_2 , $T_1 V_1^{\gamma_e - 1} = T_2 V_2^{\gamma_e - 1}$, i.e. $T_2/T_1 = (V_1/V_2)^{\gamma_e - 1}$. Then for our accretion disk problem of compression, $V_2 < V_1$, i.e. $(V_1/V_2) > 1$ and γ_e increasing from $4/3$ to $5/3$ means $(V_1/V_2)^{\gamma_e - 1}$ increases so T_2/T_1 should be larger. Fig. 5.1b shows that increases the electron adiabatic index does indeed increase the electron temperature, particularly within the inner regions where adiabatic compression is a dominant heating source.

Variables such as the definition of the jet region or the arbitrary constant s in Eq. 5.9 could also affect the electron thermodynamics. As shown in Fig. 5.1c, the jet region defined by σ_{cut} does not greatly affect the electron thermodynamics. Therefore, we keep $\sigma_{\text{cut}} = 1$ throughout these tests. In contrast, changing the constant s does affect the electron temperature since it is increasing the amount of energy removed from the gas (Fig. 5.1d). Despite the benefit of a higher s lowering the electron temperature within the ISCO, a

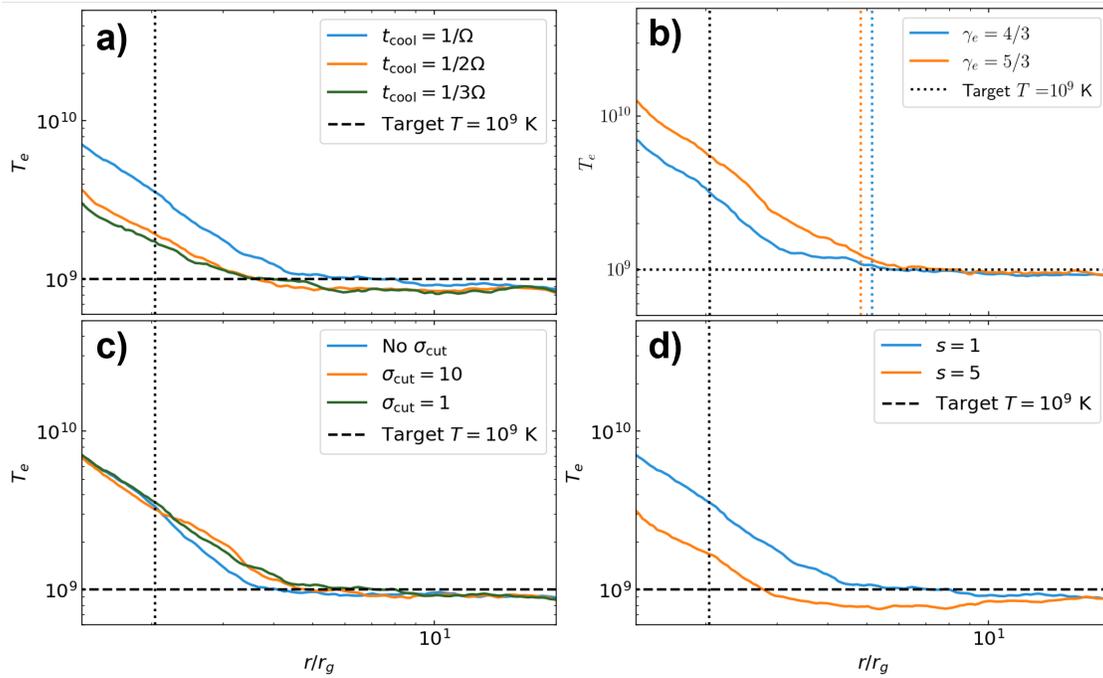


Figure 5.1: Changing a variety of parameters did not eliminate the problem of the electron temperature being much higher than the target at inner radii. The panels show the dependence of density-weighted shell-averages time-averaged from 1500 to 2500 r_g/c on a) electron cooling time, b) electron adiabatic index, c) σ_{cut} , and d) s . These parameters are defined in Eq. 5.9 or Sec. 5.2.2. Dotted vertical lines show the location of the ISCO for this black hole spin of $a = 0.9375$.

higher s also seems to overcool the gas, as can be seen by the orange line sitting below the target temperature. Because this arbitrary factor moves the cooling function away from well-approximating a derivative, we keep $s = 1$.

5.3.2 Tests Injecting Viscous Heating

The above tests are convincing evidence that the cooling function is not the problem. Next, we test the idea that the sporadic, occasionally negative viscous heating could be interacting with the cooling function in a way that leads to an electron temperature above the target. To test this idea, we use the injected viscous heating test described in Sec. 5.2.3, which is best used when Coulomb collisions are negligible. We must be careful not to inject too much viscous heating; otherwise the electron entropy floor or ceiling will trigger and lead to unphysical behavior. As such, we inject a small amount of heating, $dQ_{\text{visc}} = 10^{-6}$ in code units. Fig. 5.2 demonstrates the difference between the MHD heating described in Ch. 5.2.1 and the test injected heating. The electrons are much hotter for a simulation that injects viscous heating but does not cool electrons (blue line) compared to the MHD heating that does not cool electrons (orange line). In contrast, the electrons in the injected viscous heating test case with cooling (purple) are colder within the ISCO than the MHD heating case (green). In fact, the electrons almost exactly follow the target temperature for the injected viscosity until radii where the cooling time becomes shorter than the simulation duration, around $20r_g$. This test suggests that the electrons will not reach the target temperature due to the nature of the code's viscous heating. As such, we will simply quantify viscous heating as a source of error within the ISCO.

5.4 Extending to High Accretion Rates

The goal of this project is to extend the electron-only cooling function to the intermediate mass accretion regime where $10^{-6} \lesssim \dot{m}_{\text{Edd}} \lesssim 10^{-3}$. However, the `ebhlight` code's implementation of Coulomb collisions was never meant for high densities. In the high-density regime, explicitly solving for the Coulomb energy exchange is unstable. In particular, solving a differential equation numerically ideally involves taking small steps to approximate a derivative. High densities mean that the Coulomb exchange rate is large enough to break the assumption of small steps. In the case of Coulomb heating, we are solving the following two

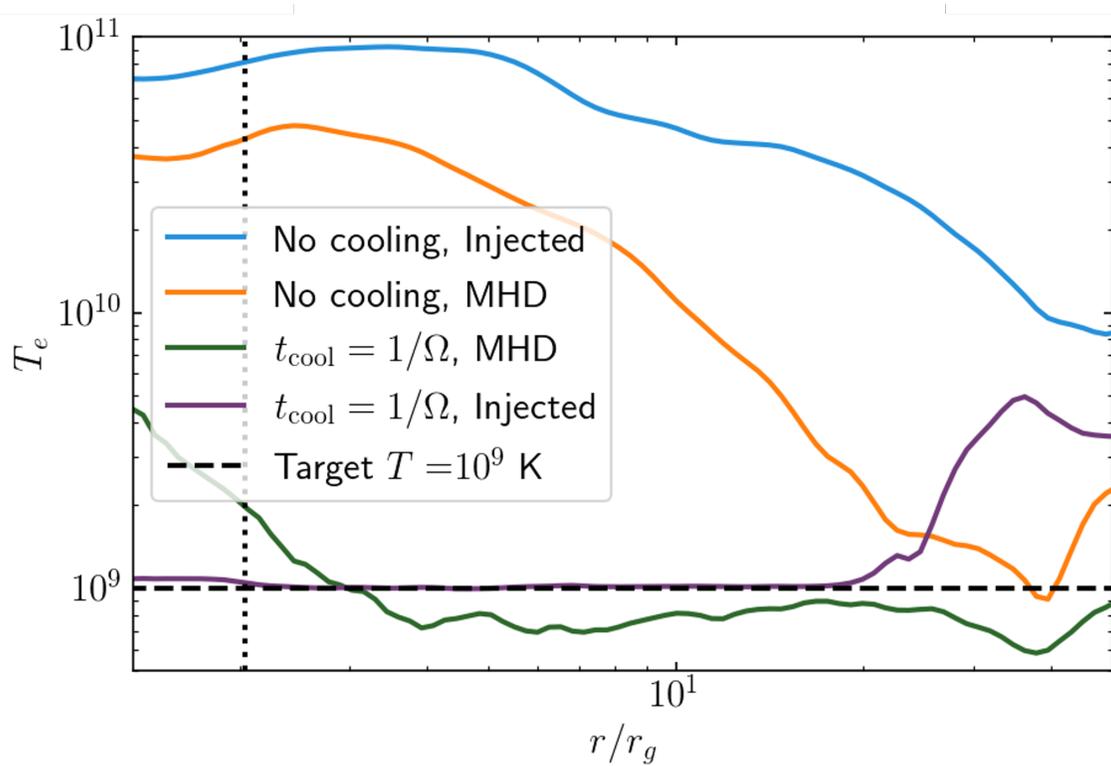


Figure 5.2: Test 2D GRMHD simulations at low accretion rate ($\dot{m} \sim 10^{-7}$) showing the impact of prescribing viscous heating as opposed to the MHD heating method. Data are density-weighted shell averages, averaged over times in the interval $(3500, 4500)r_g/c$. Dotted vertical line shows the ISCO; dashed horizontal line shows the target temperature for the electron cooling function.

equations:

$$\frac{du_e}{dt} = \frac{Q_{\text{coul}}(u_e, u_p)}{u^0} \quad (5.11)$$

$$\frac{du_p}{dt} = -\frac{du_e}{dt}, \quad (5.12)$$

where the Coulomb energy exchange rate Q_{coul} is given by Eq. 5.13.

The explicit evolution for Coulomb heating at high densities would calculate a value for the derivative large enough to completely overshoot the equilibrium solution of $T_e = T_p = T_0$. Of course, pushing the Coulomb implementation beyond the low densities it was originally intended for leads to unphysical behavior. In particular, the massive Q_{coul} would often trigger the code's ceiling on the T_p/T_e ratio and therefore fix T_p at a high value and T_e at a low value, often several orders of magnitude below the electron target temperature. In this section, I will discuss the issues I ran into when naively applying the explicit formulation of Coulomb heating to high accretion rates, and the steps I took to address their shortcomings.

5.4.1 Analytic Estimates

To get a sense for how Coulomb heating depends on number density and the temperature difference between electrons and protons, I will plot the analytic form of the Coulomb energy exchange rate from Ref. [102]:

$$Q_{\text{coul}}(T_e, T_p) = \frac{3}{2} \frac{m_p}{m_e} n_0^2 \Lambda c k_B \sigma (T_p - T_e) \left(\frac{2(\theta_e + \theta_p)^2 + 1}{\theta_e + \theta_p} \frac{K_1(\theta_m^{-1})}{K_2(\theta_e^{-1})K_2(\theta_p^{-1})} + 2 \frac{K_0(\theta_m^{-1})}{K_2(\theta_e^{-1})K_2(\theta_p^{-1})} \right), \quad (5.13)$$

where $\theta_m = 1/(1/\theta_e + 1/\theta_p)$, K_i are modified Bessel functions of the i^{th} order, and n_0 is the number density. The Coulomb logarithm Λ is fixed to 20. Note that the derivative reaches zero when $T_e = T_p$. Because the Coulomb heating rate is proportional to density squared, a factor of ten increase in density results in a factor of 100 increase in Coulomb heating.

Fig. 5.3 shows the properties of Coulomb heating rate. For a given proton temperature (assuming $T_p > T_e$), the Coulomb heating rate decreases as electron temperature increases. For a given electron temperature (assuming $T_p > T_e$), the Coulomb heating rate increases as proton temperature increases.

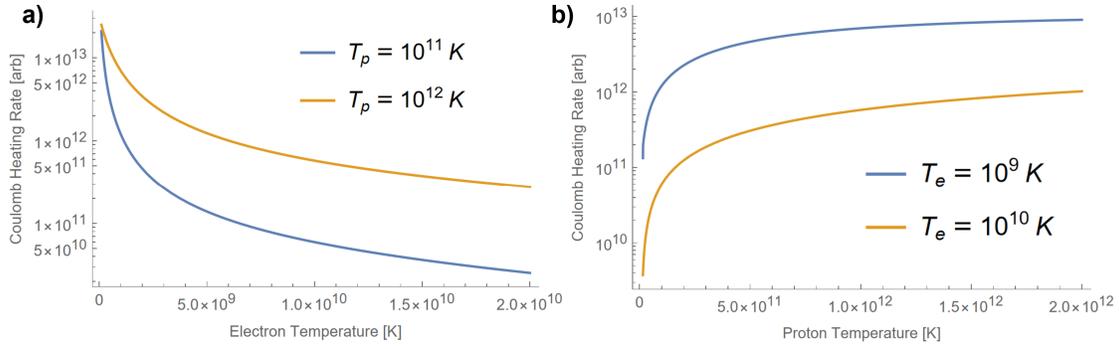


Figure 5.3: Coulomb heating rate as a function of a) electron temperature and b) proton temperature, calculated using Eq. 5.13.

Roughly, Coulomb heating rate is directly proportional to proton temperature and inversely proportional to electron temperature. The closer the temperatures get, the slower the heating rate.

Extending our simulations to the high-density regime where Coulomb heating is important risks encountering a regime where the Coulomb heating in a single simulation time step is greater than the internal energy density of the gas. To estimate whether such an error will occur in my accretion disk simulations, I took characteristic densities and temperatures and calculated the ‘‘Coulomb quality factor’’:

$$\text{Coulomb Quality Factor} = \frac{u_e/Q_{\text{coul}}}{Dt}, \quad (5.14)$$

where Dt is the simulation timestep. A Coulomb quality factor of $\lesssim 1$ or even 10 will likely result in unphysical behavior, since it is not well approximating a derivative. In Fig. 5.4, I set $Dt = 6 \times 10^{-3}$ as an estimate pulled from the output of high accretion rate simulations. For the low accretion rate (Fig. 5.4a), the simulation timestep is never an issue. For the large accretion rate (Fig. 5.4b), there is an interesting dependence on electron and proton temperature. When the protons are about the same temperature as the electrons but the electrons are hot (5×10^{10} K), the simulation timestep is very close to the Coulomb timestep. These estimates suggest that cooling down the gas gradually might prevent unphysical behavior even in the large accretion rate case.

Fig. 5.4c shows the specific case of electrons at the target temperature $T_{e,\text{target}} = 10^9$ K. From this plot it is clear that protons need to be cooler than about 10^{10} K for the simulation timestep to be short enough at high densities experienced in the simulation.

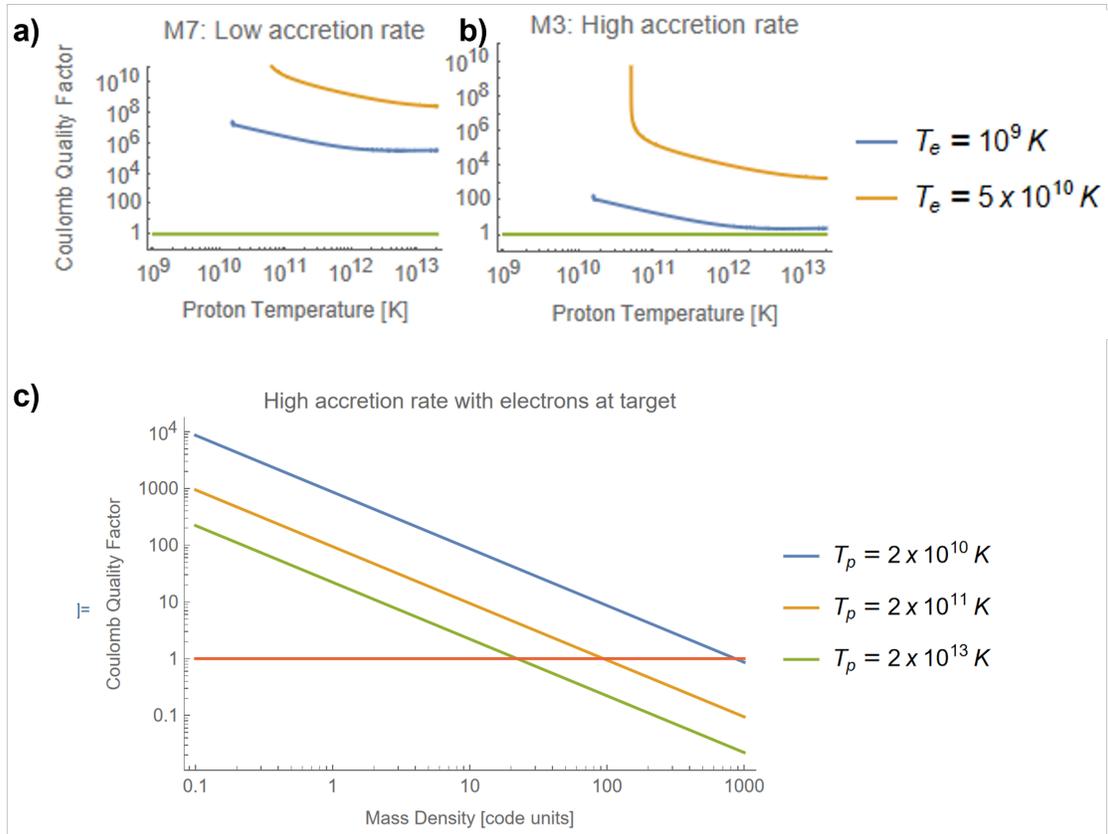


Figure 5.4: Plots of the Coulomb quality factor (Eq. 5.14) at low accretion rates (panel a) and high accretion rates (panel b) over proton temperature. Panel c fixes the electron temperature at a target value of $10^9 K$ and shows how the Coulomb quality factor depends on mass density.

5.4.2 Numerical Attempts with Explicit Coulomb Heating

Although we already know that `ebhlight` was never meant for higher accretion rates, estimates in Ch. 5.4.1 suggest that if I can slow down cooling enough and approach the equilibrium solution slowly enough, then perhaps the code will not crash. These approaches ultimately did not succeed and led to the need to implement an implicit Coulomb solver, outlined in Ch. 5.4.3.

5.4.2.1 Empirical Markers of the problem

As described above, the large values for Q_{coul} lead to reversals in the ordering of T_e and T_p that would not occur if the derivative were well-approximated because it would reach zero. This error manifests in several different diagnostics, beginning in Q_{coul} itself, moving to T_e/T_p and T_e , and finally affecting the large-scale dynamics of the accretion disk because of the reaction on the total gas. The Coulomb quality factor (Eq. 5.14) often provides even more evidence for this problem. Fig. 5.5 shows the difference between high and low accretion rates for 3D simulations with resolution 320x256x160. In the problematic regions, the electron temperature dips an order of magnitude below the target temperature (a), the ratio T_p/T_e hits its ceiling value (b), the scale height-to-radius ratio H/r becomes wiggly (c), and the Coulomb quality factor becomes less than 1 (d).

To further diagnose the problem, vertical slices taken at $\phi = 0$ at a given time are shown in Fig. 5.6. The top row shows that the electron temperature drops below the target in the midplane of the accretion disk, where densities are higher and Coulomb coupling is stronger. The high T_p/T_e is driven by small electron temperatures in the midplane. The energy exchange rates in the bottom row show how the magnitudes compare. Viscous heating is sporadic but generally averages to a positive value. Electron cooling stops occurring in the disk midplane around $20r_g$, presumably because the electron temperature has fallen below the target. The last panel, with Coulomb energy exchange, shows that Q_{coul} is negative in the midplane. A negative value for Q_{coul} means that electrons are hotter than protons, which contradicts the output of T_p/T_e . This figure demonstrates that the Coulomb overshoot happens within a single timestep.

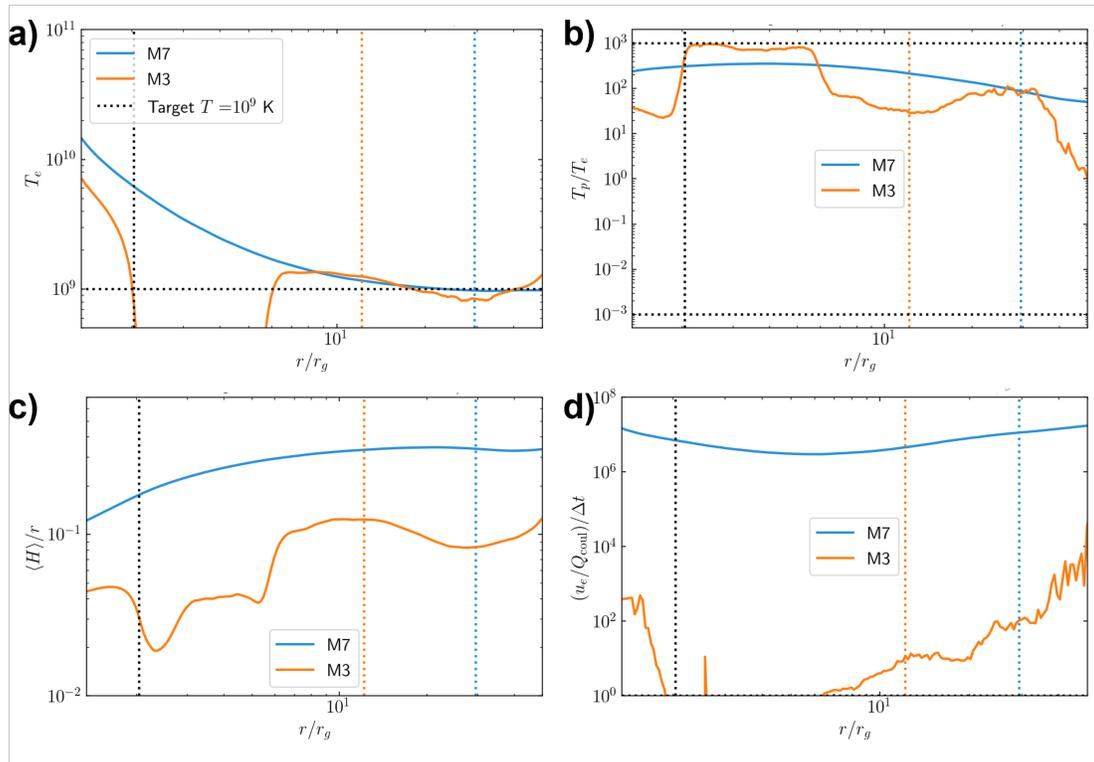


Figure 5.5: Illustration of unphysical behavior due to explicit evolution of Coulomb heating in the high-density regime. Plots show density-weighted shell averages taken from high-resolution GRMHD simulations, averaged over $13,000$ to $15,000 r_g/c$.

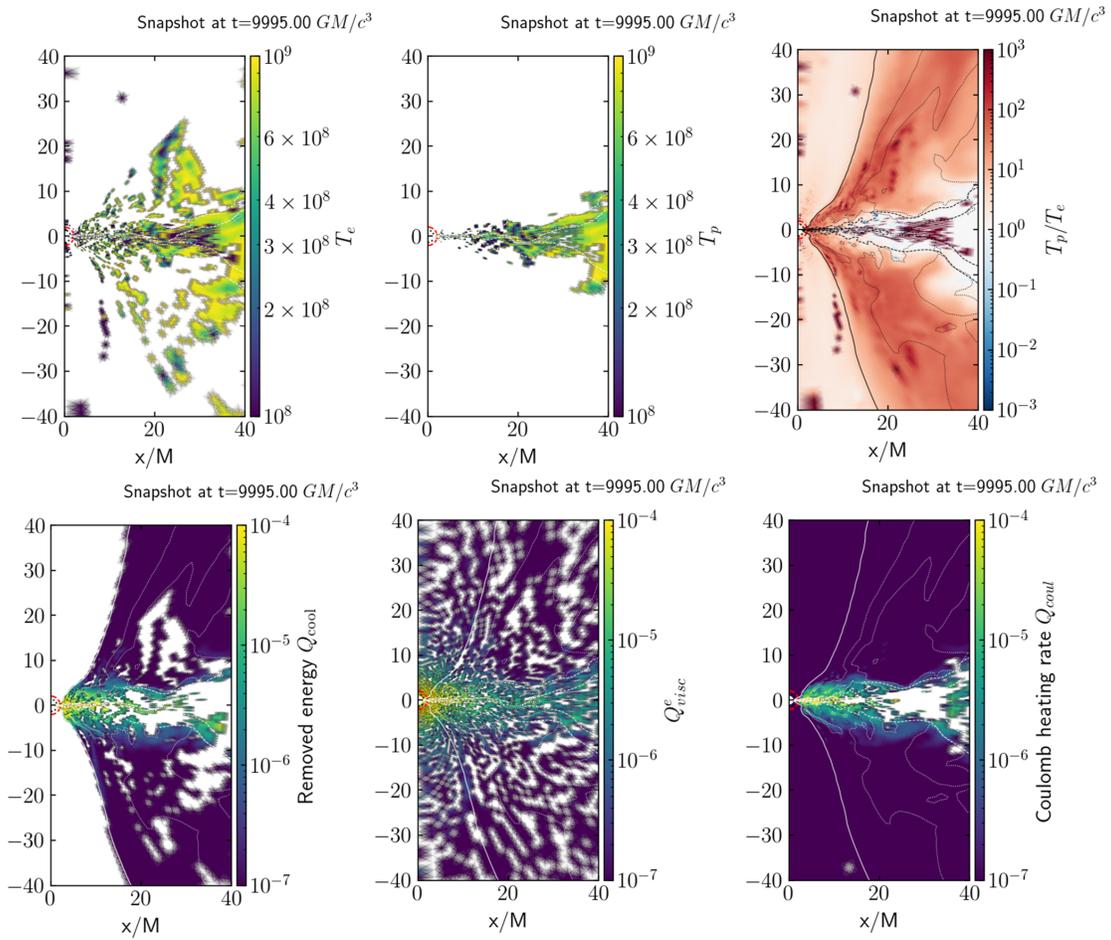


Figure 5.6: The unphysical Coulomb exchange rate that breaks the explicit evolution happens predominantly in the midplane of the accretion disk, where densities are highest. These slices are taken at $\phi = 0$ at a snapshot in time from high-resolution 3D GRMHD simulations.

5.4.2.2 Attempted Solutions

Before moving to the best solution of an implicit Coulomb solver (see Ch. 5.4.3), many other attempts to continue using the explicit evolution were made. Based off the idea that cooling and Coulomb exchange generally needed to be slowed down (Ch. 5.4.1), we attempted to step closer and closer to the final equilibrium state rather than start far from equilibrium and hope the algorithm held. For these attempts, we generally restarted from a simulation with smaller accretion rate and changed variables such as the target temperature and the cooling time. We also tried implementing a Coulomb limiter that would trigger whenever calling the Coulomb function resulted in the ordering T_e and T_p reversing. Finally, we thought that implementing isothermal electrons might fix the problem; see Ch. 5.C. However, none of these attempts at fixing the explicit evolution led anywhere. Instead, we move on to implementing the implicit Coulomb evolution.

5.4.3 Implementing an Implicit Coulomb Solver

The way `ebhlight` implements viscous and Coulomb heating is with operator-splitting, which uses the same fluid variables to calculate the values of source terms (viscous/Coulomb heating) and then adds those together. The update of the final values u_e^f then reads

$$u_{e,f}^f = u_{e,i}^f + Q_{\text{coul}}(u_e^s) \frac{\Delta t}{u^0}, \quad (5.15)$$

where the subscripts indicate the array (Pf or Ps). The explicit method simply sets Q_{coul} using $u_{e,i}^s$, i.e. just takes the values from the previous iteration and sets the Coulomb energy exchange rate.

To extend `ebhlight` to larger accretion rates, we need to implicitly solve for the Coulomb energy exchange. We will therefore solve for $u_{e,f}^s$. From here on, we will drop the superscripts since they will all be s ; all that follows will serve to update u_e^f as in Eq. 5.15. Because total internal energy density u_g is conserved, we can condense the system of equations in Eq. 5.11 and 5.12 into a single equation using the constraint of Eq. 5.1. The final equation reads:

$$\frac{du_e}{dt} = \frac{Q_{\text{coul}}(u_e, u_g - u_e)}{u^0}. \quad (5.16)$$

We discretize this equation with the Crank-Nicolson method [182] that takes the internal energy density u_e^n

at step n to u_e^{n+1} at step $n + 1$.

$$\frac{u_e^{n+1} - u_e^n}{\Delta t} = \frac{1}{2} \frac{1}{u^0} \left[Q_{\text{coul}}(u_e^{n+1}, u_g - u_e^{n+1}) + Q_{\text{coul}}(u_e^n, u_g - u_e^n) \right]. \quad (5.17)$$

Because the right hand side of this equation is highly nonlinear in u_e (see Eq. 5.13), we will solve it numerically using Brent's method [182]. Brent's method requires an interval $[a, b]$ over which the right-hand side switches side. With this problem we have an easy choice for the interval: assuming $T_e < T_p$ initially, the minimum for T_e 's solution is T_e and the maximum is the equilibrium temperature T_0 (see Appendix 5.A). In `ebhlight`, we could choose to use one of two arrays in calculating the values of Eq. 5.17. We choose to maintain operator splitting, and therefore set the source terms and the difference $\delta u_e = u_e^{n+1} - u_e^n$ using values from the `Ps` arrays, which are then used to update the `Pf` array's u_e .

To avoid solving the equations in the regime where we know ahead of time what the solution will be, we avoid the implicit solver in the single-temperature regime. Specifically, if $|T_e - T_0|/T_0 < \epsilon$, $|T_p - T_0|/T_0 < \epsilon$, and the calculated change in internal energy is large enough that $10\Delta u_e > u_e$, the solver will simply set $T_e = T_p = T_0$ instead of implicitly solving Eq. 5.17.

5.4.3.1 Gas Box Implementation Tests

For a well-controlled set-up, I test the implementation of the implicit Coulomb heating in a 2D, doubly-periodic box with a Minkowski metric. This set-up eliminates any viscous heating and adiabatic heating/cooling. With this set-up, I can look at individual time steps to examine the numerical method.

As a first example, consider the high-density regime and start with electrons at 10^8 K and protons at 2×10^9 K. Without cooling (i.e. by setting the target temperature to a large value of 10^{12} K), the two species should equilibrate at a temperature between 10^8 and 2×10^9 K. As seen in Fig. 5.7, the explicit solver breaks down in this regime — as expected, because it was never meant for this regime. The implicit solver, on the other hand, quickly equilibrates and maintains the same temperature for both species.

We also need to compare the operator-split vs. not operator-split implementation of the implicit Coulomb solver described in Ch. 5.4.3. In this set-up, the electrons and protons start with the same temperature $T_e = T_p = 10^{10}$ K, sitting above the target temperature of 10^9 K. When the simulation begins,

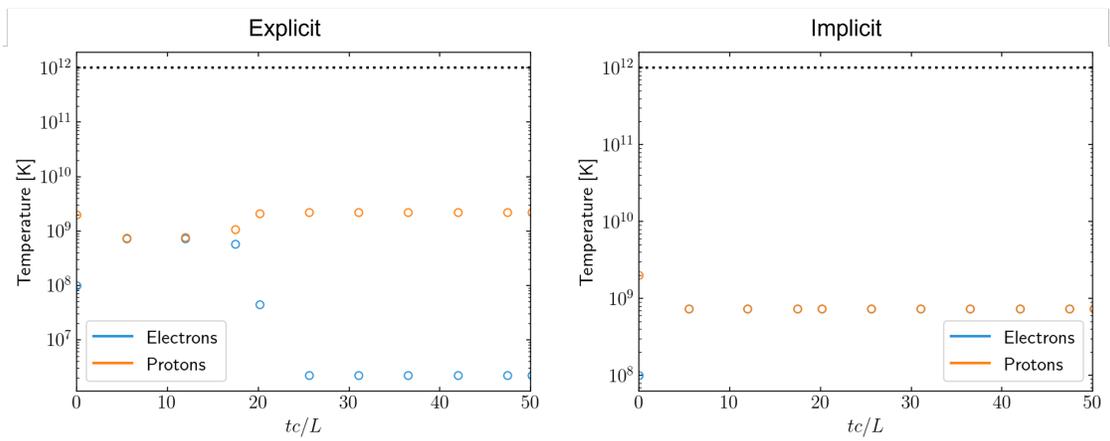


Figure 5.7: Gas box test for Coulomb thermalization in the high-density, single-temperature regime where the explicit solver (left) breaks and sets T_p/T_e to its ceiling value, compared to the implicit solver (right) that is meant for high densities.

the two species should stay in the single-temperature regime and cool down to the target temperature. However, as seen in the left panel of Fig. 5.8, the operator-split method breaks down. The non-operator-split method that uses Pf values to evaluate Coulomb heating behaves as physically expected. This difference in the single-temperature regime motivates our use of the non-operator-split regime in subsequent tests.

We outline several more tests in Fig. 5.9 that all use the non-operator-split implicit Coulomb heating. The first test (shown in panel a) investigates the Coulomb energy exchange without any interference from the electron cooling function. The electron and protons evolve smoothly towards their equilibrium temperature. The second test (panel b) shows the case of electron cooling in a very low density case such that Coulomb collisions are negligible. As expected, protons cannot cool except via Coulomb collisions and thus remain at high temperatures, whereas the electrons cool down to the target temperature. In the third test, the density is high enough to thermalize protons and electrons all the time, such that when electrons cool down to the target temperature, protons follow closely. Finally, the combination of Coulomb collisions and electron cooling demonstrates how the electrons initially heat up due to a large temperature difference, then cool down as the temperature difference decreases (see Fig. 5.3). All of these tests break the explicit solver due to the high densities.

5.4.3.2 2D Torus Implementation Tests

Next, I show the comparison of the three methods for Coulomb heating in 2D GRMHD models of a thick accretion disk. These tests are 2D Fishbone-Moncrief tori with an inner radius of $6r_g$ and a resolution of 160×128 cells in the r and θ directions, respectively. These simulations have a weak magnetic field. To ensure that the implicit Coulomb coupling does not affect the gas dynamics, I first show the evolution in the low accretion rate case where electrons and protons do not interact strongly. The evolution stays relatively similar for all implementations, although the operator split implicit method heats up electrons more by about a factor of 2 within the ISCO. The ratio $Q_{\text{coul}}/Q_{\text{visc}}^e$ stays below one at all radii, indicating the subdominance of Coulomb heating in this regime.

At high accretion rate $\dot{m} \sim 10^4 \dot{m}_0$, high densities begin to bring parts of the accretion disk into a regime where Coulomb collisions cannot be treated explicitly. The Coulomb quality factor (Eq. 5.14)

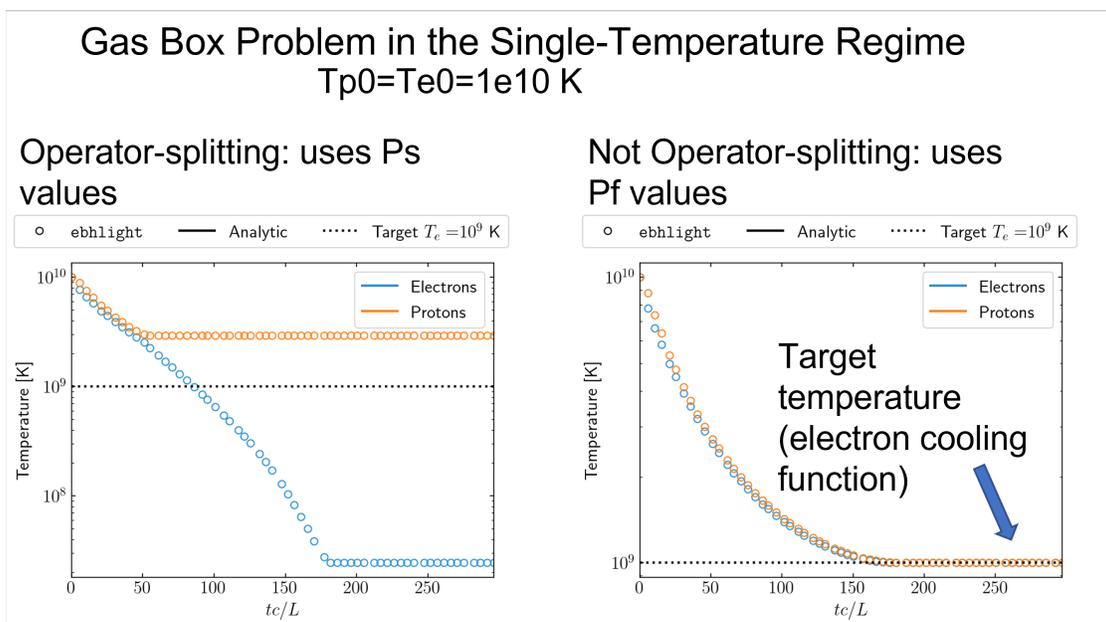


Figure 5.8: Gas box test for Coulomb heating in the single-temperature regime. The electrons and protons should cool down to the same temperature to the target, but the operator-splitting implementation breaks down (left) whereas the non-operator-splitting implementation behaves as expected (right).

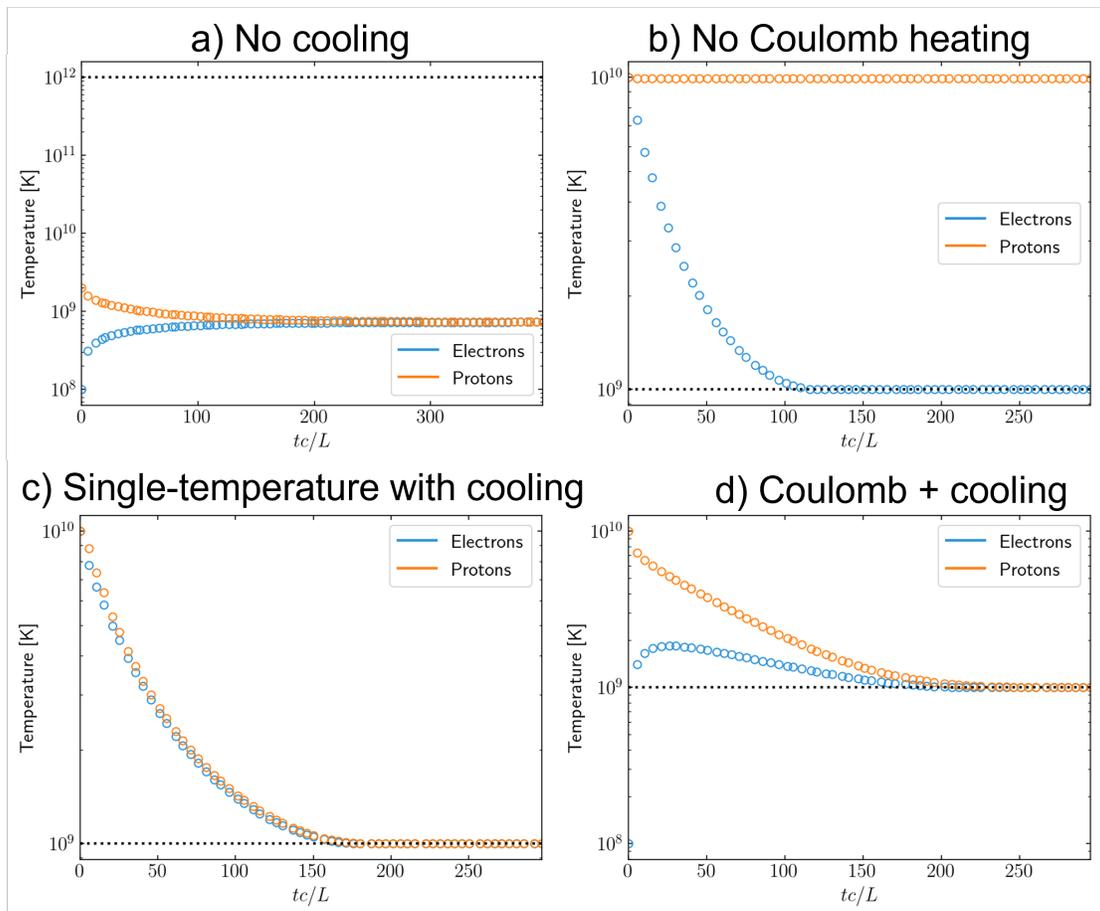


Figure 5.9: Two-species thermalization tests for gas box in flat space-time.

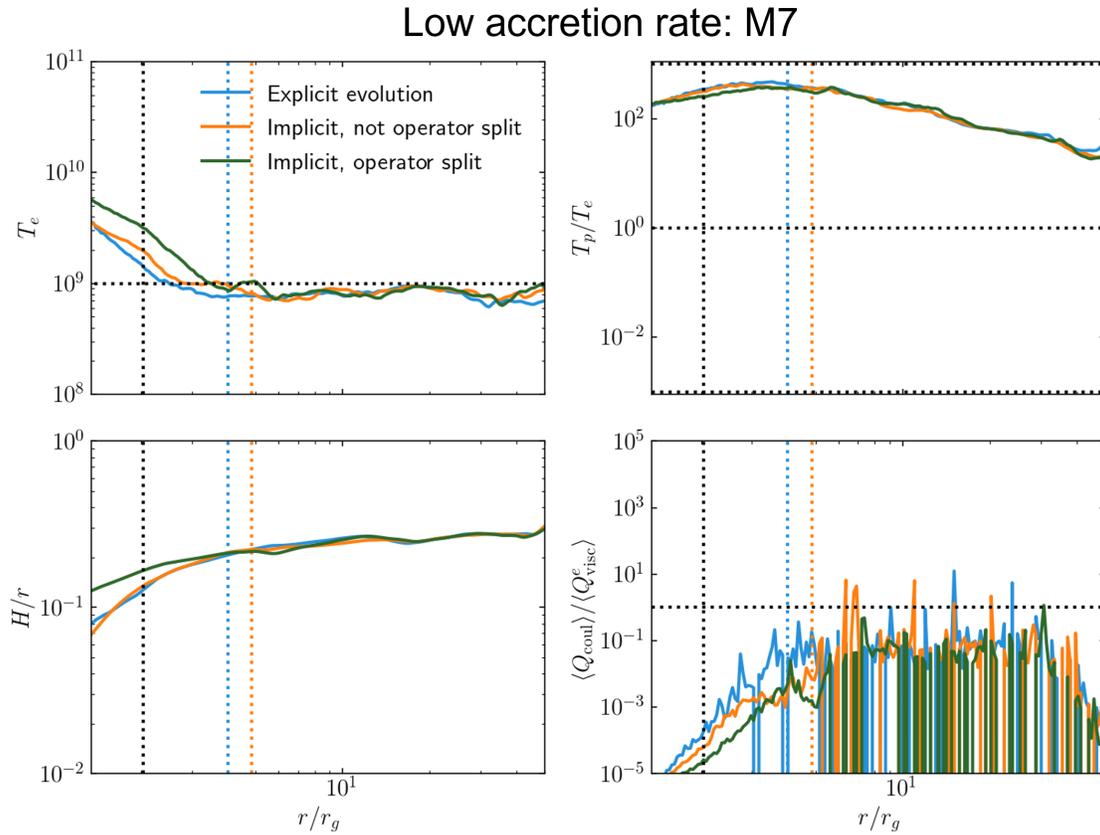


Figure 5.10: 2D GRMHD tests of the Coulomb heating implementation at low accretion rate M7, where it should not make a difference to use an implicit or explicit evolution. Plots show density-weighted shell averages time-averaged from $4500 < tc/r_g < 5000$. Blue lines show the explicit evolution, which holds for this low density limit, while orange and green show the two implicit methods, one that uses operator-splitting (green) and one that does not (orange). Compare to the high-density regime in Fig. 5.11.

becomes less than 1 in some places, resulting in spikes in the ratio T_p/T_e for the explicit evolution as the floor and ceiling values for this ratio kick in. As parts of the disk enter the single-temperature regime, the operator split implicit method also begins to break down, seen in the small spikes of T_p/T_e in the green line. Only the non-operator split implicit method maintains a smooth temperature ratio at the radii where the disk begins to collapse. Because the Coulomb energy exchange rate Q_{coul} is effectively smaller in the implicit implementation compared to the (unlimited) explicit implementation, the electron temperature is higher within the ISCO for the implicit method. The higher electron temperature and higher ratio T_p/T_e also leads to a thicker disk within the ISCO. The ratio of Coulomb to viscous heating crosses 1 right around the ISCO, demonstrating that Coulomb collisions dominate outside the ISCO.

5.5 Results

Because implementation and testing of the implicit Coulomb solver is ongoing, all results presented in this section are tentative. In particular, the 3D simulations do not have production-run resolution and therefore do not resolve the MRI. Because the MRI is an important source of viscous dissipation in these weakly-magnetized disks, Q_{visc} is likely too small, which could artificially lead to collapse of the disk. These simulations have a resolution of (160, 128, 80) cells in the (r, θ, ϕ) directions, respectively. The higher accretion rate simulations are restarted from the lowest accretion rate (M7) at a time of $7000 r_g/c$.

5.5.1 Temporal Evolution

As seen in Fig. 5.12, the M3 simulation is not in steady-state: the proton temperature is decreasing rapidly at both radii shown. The higher accretion rate simulation M1 reaches an equilibrium from the outside-in. M1's proton temperature at large radius $10r_g$ reaches the electron temperature shortly after restarting, around $7500 r_g/c$. At the ISCO, M1's proton temperature reaches the electron temperature only at $12500 r_g/c$, $5500r_g/c$ after restarting. The mid-accretion rate M3 simulation has not reached equilibrium, but is evolving quickly.

The disk collapsing from the outside-in might run counter to naive expectations that the disk should evolve from the inside-out, since the dynamical time is much shorter at smaller radii. Fig. 5.13 shows relevant

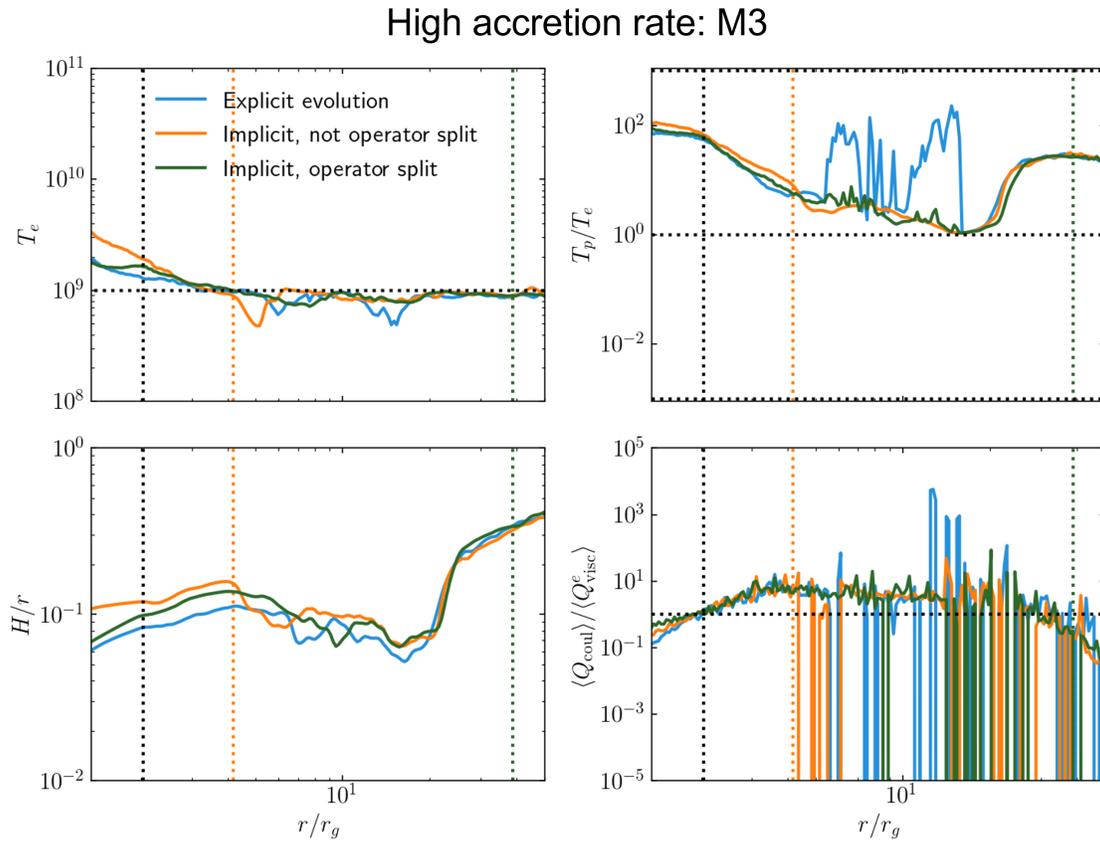


Figure 5.11: 2D GRMHD tests of the Coulomb heating implementation at high accretion rate M3; density-weighted shell averages time-averaged from $4500 < tc/r_g < 5000$. Blue lines show the explicit evolution, which breaks down for this high density limit, as seen in particular in the ratio T_p/T_e . Orange and green lines show the two different implicit method implementations. Because the operator split method (green) shows signs of breaking down in the single-temperature limit, we prefer the non-operator split method (orange). Compare to the low-density regime in Fig. 5.10.

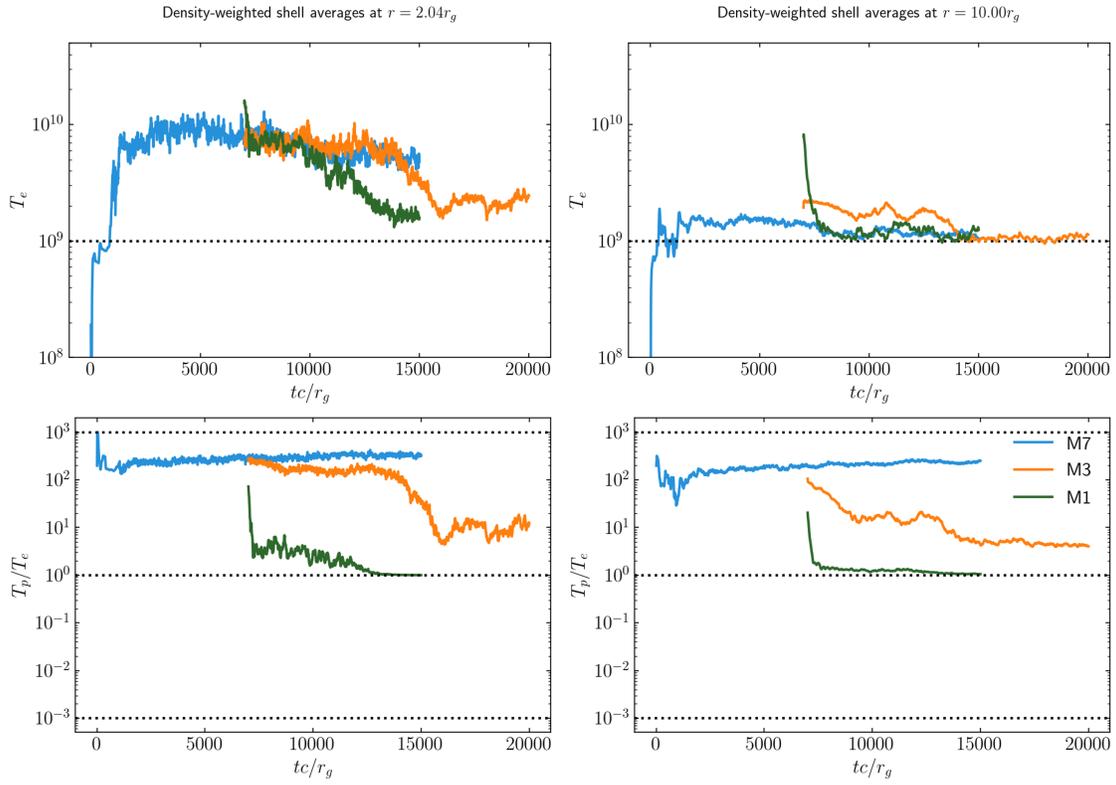


Figure 5.12: Density-weighted shell averages at $r = 2.04r_g$ (left) and $r = 10.0r_g$ (right) plotted over time to demonstrate the lack of steady-state. The speed of equilibration depends both on the radial location and the accretion rate. The target temperature 10^9 K is shown in the top row as a dotted horizontal line. Dotted lines in the bottom row show the minimum and maximum ratios for T_p/T_e and the single-temperature case of $T_p/T_e = 1$.

timescales for protons and electrons, defined as follows:

$$\begin{aligned}
 t_{\text{heat}}^i &= \frac{u_p}{Q_{\text{visc}}^i} & t_{\text{heat}}^e &= \frac{u_e}{Q_{\text{visc}}^e} \\
 t_{\text{Coul}}^i &= \frac{u_p}{Q_{\text{coul}}} & t_{\text{Coul}}^e &= \frac{u_e}{Q_{\text{coul}}} \\
 t_{\text{cool}} &= \frac{1}{\Omega} & t_{\text{cool}}^e &= \frac{u_e}{Q_{\text{coo}}}.
 \end{aligned} \tag{5.18}$$

Finally, $t_{\text{infall}}(r) = \int_r^\infty u^0 dr$.

In absolute terms, t_{Coul}^e is longer at larger radii. However, relative to the infall time, t_{Coul}^e is much shorter at $10r_g$ and close to the infall time at the ISCO. For M7, the Coulomb time is greater than the infall time for both electrons and protons for $r \lesssim 40r_g$. For M1 and M3, the Coulomb time is shorter than the infall time until within the ISCO, where the infall time becomes shorter than the Coulomb time. For M1 and M3, the electron cooling and electron Coulomb time are practically identical, showing that these two physical processes balance each other.

5.5.2 Radial Structure

Ideally, we would average over an infall time after the simulations reached steady-state. However, time constraints limited the amount of time these simulations were run for. Given this caveat, it is nevertheless interesting to examine the radial structure of the accretion disks as a function of accretion rate. Fig. 5.14 shows the electron temperature, T_p/T_e , H/r , and $Q_{\text{coul}}/Q_{\text{visc}}^e$ shell-averages as a function of radius. As noted in Ch. 5.2.3, the electron temperature does not reach the target temperature in the inner portions of the accretion flow. The lowest accretion rate, M7, has protons that are orders of magnitude hotter than the electrons everywhere within the inner $50r_g$. In contrast, the protons in the M3 run have begun to cool down to the electron temperature. The highest coupling results in a disk at the single-temperature regime, seen by $T_p = T_e$ for the M1 simulation. Both M3 and M1 appear to be cooling towards the thin disk H/r scaling. The final panel shows that for M7 Coulomb collisions are always subdominant to viscous heating in that $Q_{\text{coul}}/Q_{\text{visc}}^e < 1$. For M1 on the other hand, $Q_{\text{coul}}/Q_{\text{visc}}^e > 1$ everywhere, showing that Coulomb collisions dominate viscous heating. The simulation M3 experiences both regimes, where Coulomb collisions are subdominant at $r \lesssim r_{\text{ISCO}}$ and dominant for $r \gtrsim r_{\text{ISCO}}$. It is this transition that could lead to disk truncation.

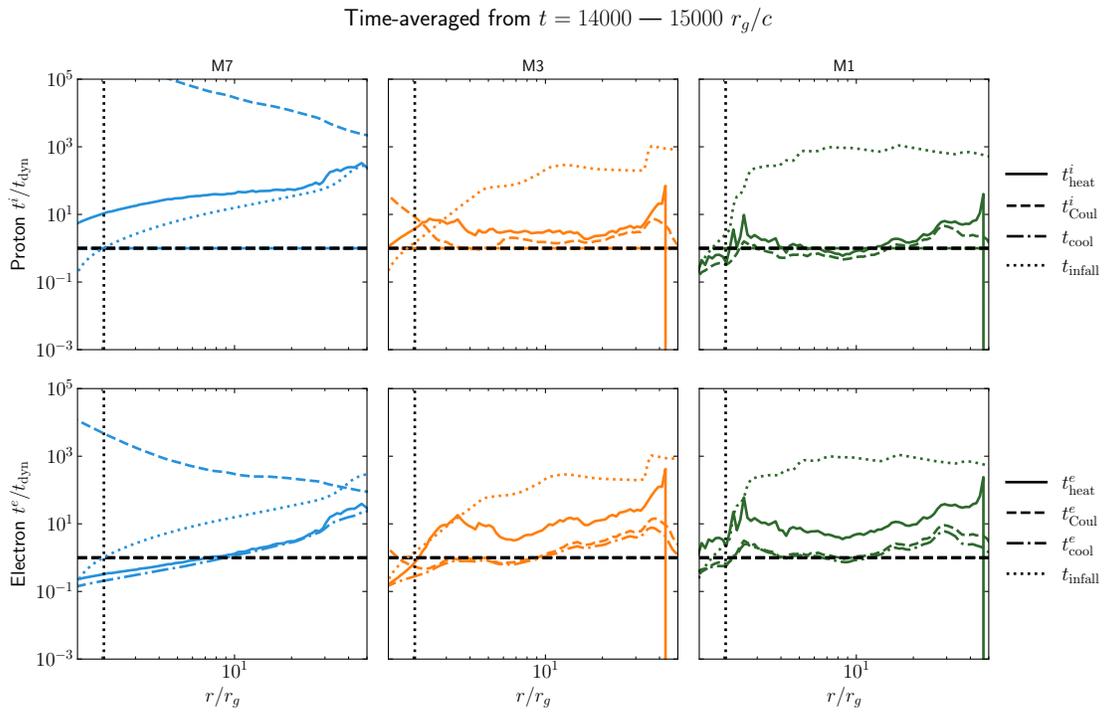


Figure 5.13: Density-weighted shell averages of relevant proton and electron timescales for three different accretion rates. Timescales are normalized to the dynamical time $t_{\text{dyn}} = 1/\Omega$. Here, the black vertical dotted line shows the location of the ISCO. The black horizontal dotted lines show where the timescales equal the dynamical time.

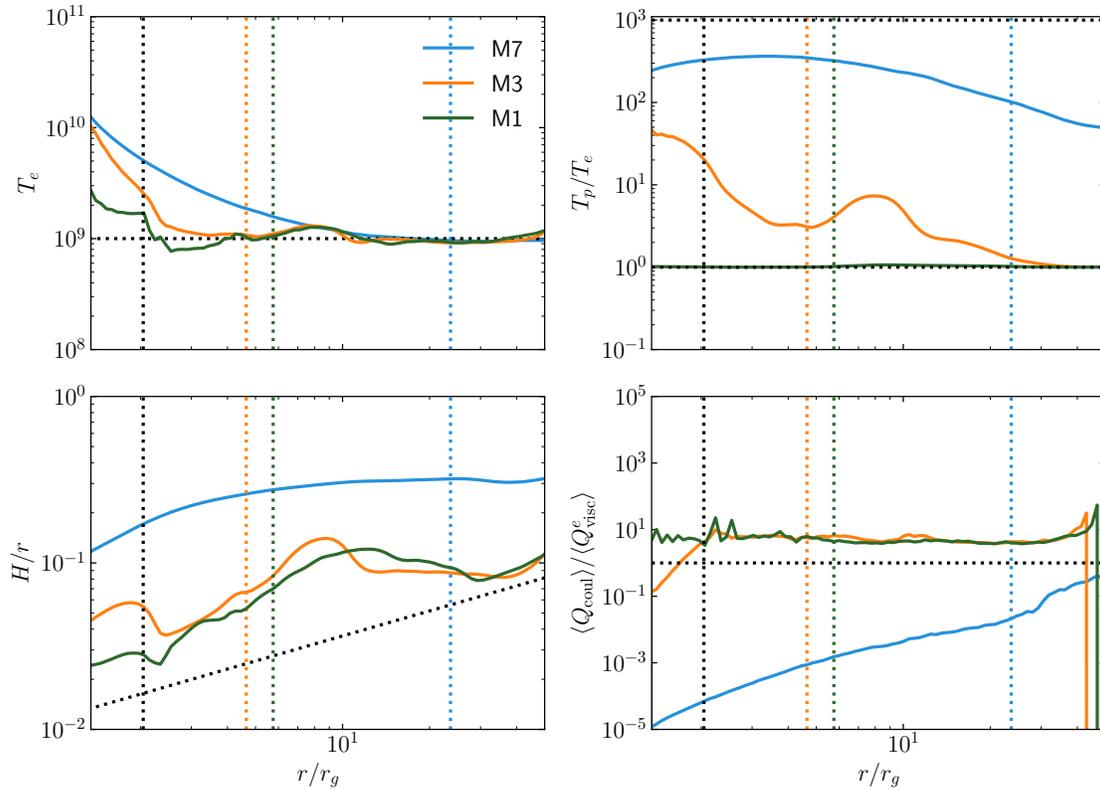


Figure 5.14: Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. Dotted vertical black line shows the location of the ISCO. Dotted vertical colored lines show the location of the infall equilibrium radius r_{eq} where $t = r_{eq}/|v^r|$. Dotted black lines in the T_p/T_e and $Q_{\text{cool}}/Q_{\text{visc}}^e$ plots shows where the ratio equals 1. Dotted black line in the H/r plot shows the theoretical thin disk scaling for $T_p = T_e = T_{\text{target}}$.

More diagnostics are plotted in Fig. 5.15. The first panel shows that the Coulomb quality time (Eq. 5.14) is much greater than 1 for all simulations, showing that the implicit Coulomb implementation successfully avoids overcooling via Coulomb collisions. The second panel shows that $Q_{\text{cool}}/Q_{\text{cool}} \sim 1$ for M1 and M3, indicating that these two simulations are in the regime where Coulomb collisions balance cooling. That is in contrast to M7, where cooling instead balances viscous heating rather than Coulomb heating. The bottom row shows the MRI quality factors to demonstrate that the simulations are rather severely underresolved, leading to potential problems.

5.5.3 Vertical Structure

To get a sense for the vertical structure of the simulations, we plot phi-averaged slices, time-averaged over the same time period as the density-weighted shell averages in Fig. 5.16. The higher accretion rate simulations M3 and M1 become strongly magnetized in the inner disk regions, reaching $\beta < 1$ particularly within $r \lesssim 25r_g$. Fingers of less-strongly magnetized gas reach all the way down to the event horizon. The density, shown in code units, clearly increases and forms a dense core for M3 and M1. The highest accretion rate simulation M1 has a denser outline surrounding the main body. The ratio T_p/T_e clearly shows how Coulomb collisions bring the electrons and protons into thermal equilibrium. For M7, protons are always much hotter than electrons. For M3 and M1, the body of the disk has protons and electrons at the same temperature. The highest accretion rate has a noticeably larger area where $T_p/T_e \sim 1$, extending to larger heights above the midplane. The final panel shows how Coulomb collisions dominate the electron heating within the disk body, whereas viscous heating dominates at larger heights.

5.5.4 Higher Target Temperature

A constant electron temperature of 10^9 K would mean a disk scale height-to-radius ratio of 0.018 at the ISCO. Such a small H/r pushes the limits of computational capabilities to resolve the MRI. As such, we also examine a target temperature of 10^{10} K, which gives $H/r = 0.057$ at the ISCO. It is currently unclear how much of an impact this change in target temperature will have on the overall physics of the disk, since 10^{10} K could be reaching into a physically different regime.

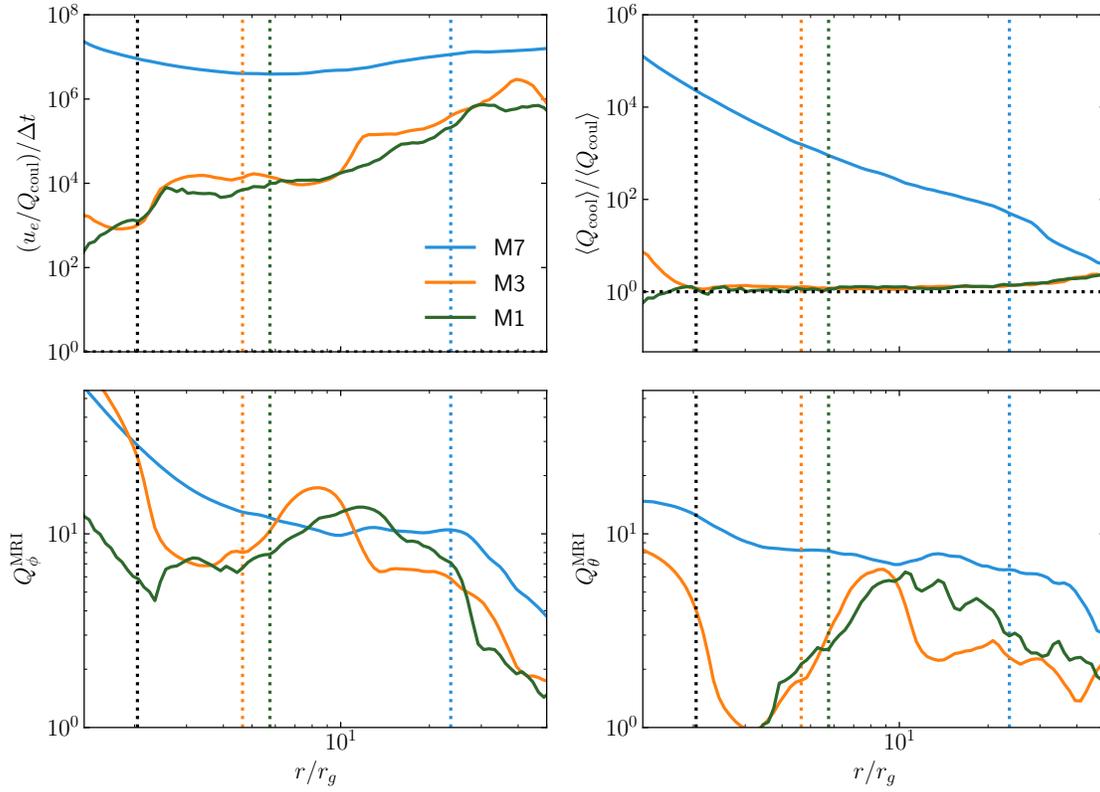


Figure 5.15: Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. Dotted vertical black line shows the location of the ISCO. Dotted vertical colored lines show the location of the infall equilibrium radius r_{eq} where $t = r_{eq}/|v^r|$.

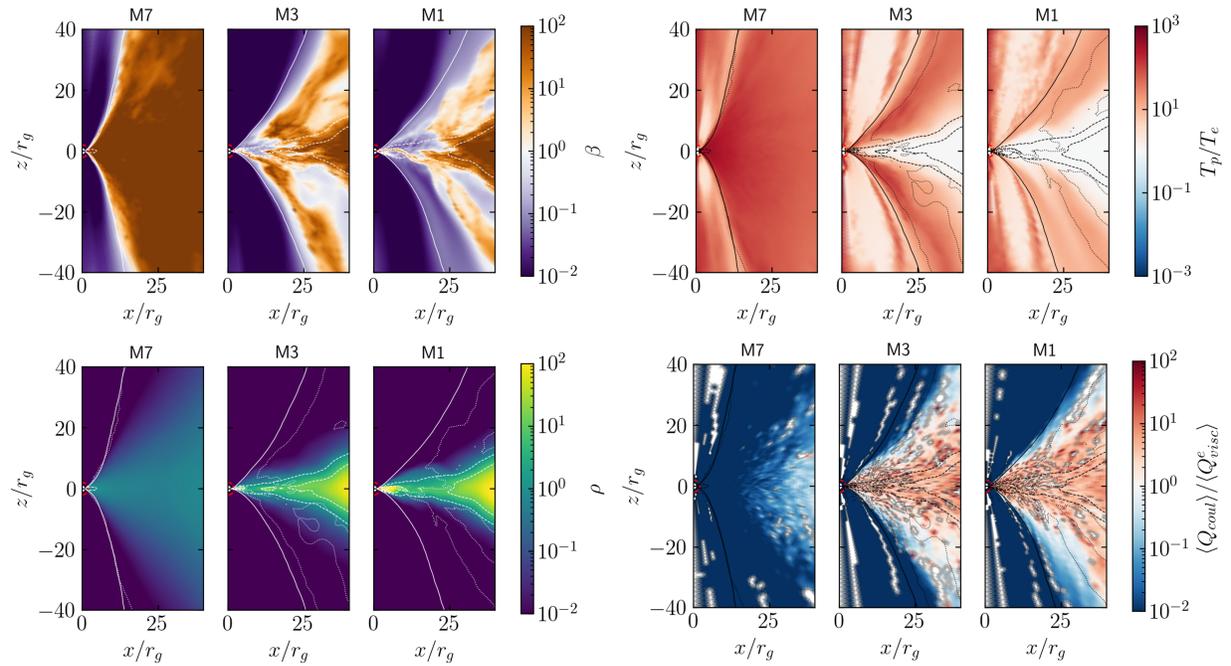


Figure 5.16: Azimuthally-averaged vertical slices for three accretion rates, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. The solid contour shows $\sigma = 1$. The dashed contours show $\rho = 1$ and 10 . The dotted contours show $\beta = 1$.

The results are shown in Fig. 5.17. Note that the simulation with a target temperature of 10^{10} K shows that T_p is dropping over the course of this time-average, and therefore has not reached steady state yet. The electron temperature is obviously higher for the larger target temperature, but T_p/T_e is approximately the same. The protons therefore have a higher temperature in the higher target temperature case. Neither simulation has reached its target H/r , again probably due to the need for a longer evolution period. The ratio $Q_{\text{coul}}/Q_{\text{visc}}^e$, however, looks similar.

5.5.5 Constant H/r

In the thin disk regime, we might expect a constant H/r rather than a constant target temperature that was motivated by optically-thin inverse Compton scattering. In this section, we explore preliminary results for simulations where the target temperature was set to scale as $T \propto 1/r$.

We compare the results in Fig. 5.18; note these are not yet in a steady state. As before, the higher accretion rate simulation M1 has electron temperatures close to the target even within the ISCO. The M1 simulation's proton-electron temperature ratio is only about a factor of 3 greater than the target, compared to M3's factor of 10 or more, though that could be due to M3 taking a longer time to equilibrate. The values for H/r are not close to their targets (dotted lines), which could be a result of low spatial resolution or an indication that steady-state has not yet been reached.

5.6 Discussion and Conclusions

Although this work is still in the early stages, we have demonstrated the potential of an electron-only cooling function. After implementing an implicit solver for Coulomb energy exchange, we show that at high accretion rates $\dot{m} \sim 0.1\dot{m}_{\text{Edd}}$, an initially hot accretion flow quickly approaches the single-temperature regime. We have also found preliminary evidence for a puffy region within the ISCO at intermediate accretion rates M3, which needs to be further explored. In the future, I aim to determine the persistence of this transition and to examine its dependence on accretion rate in high-resolution simulations.

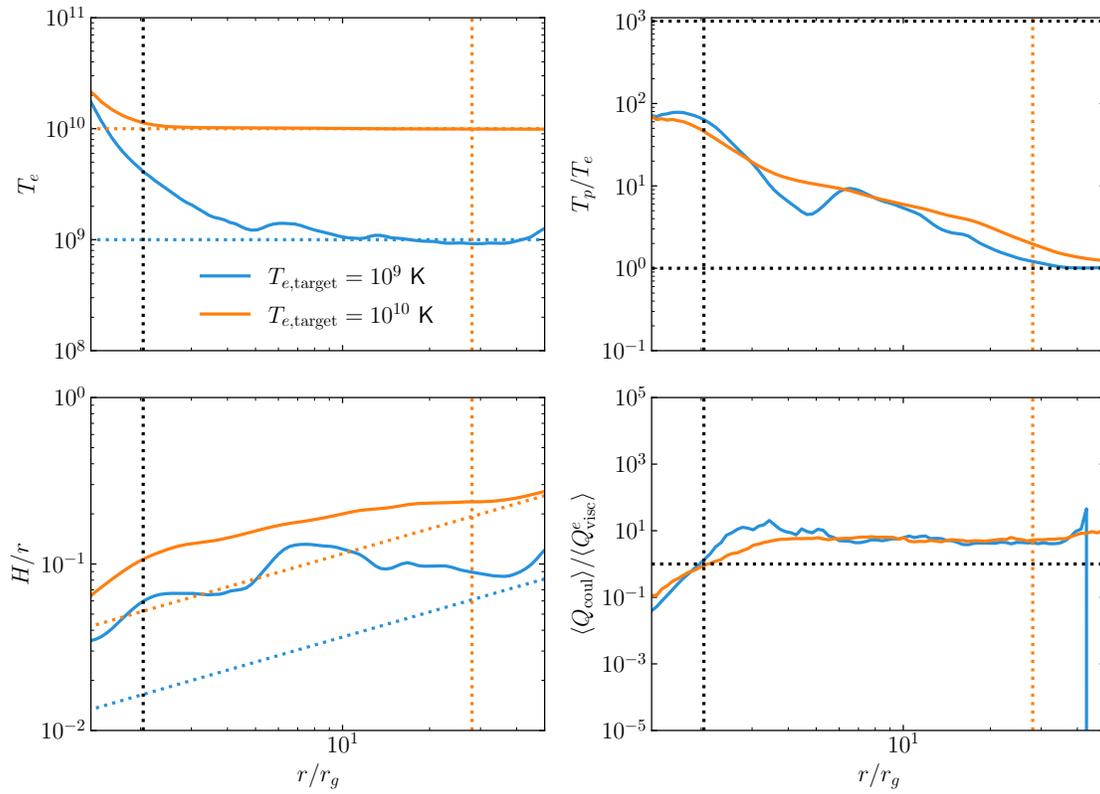


Figure 5.17: Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$. Simulations have an accretion rate that is about the same, $\dot{m} \sim 10^{-3} \dot{m}_{\text{Edd}}$.

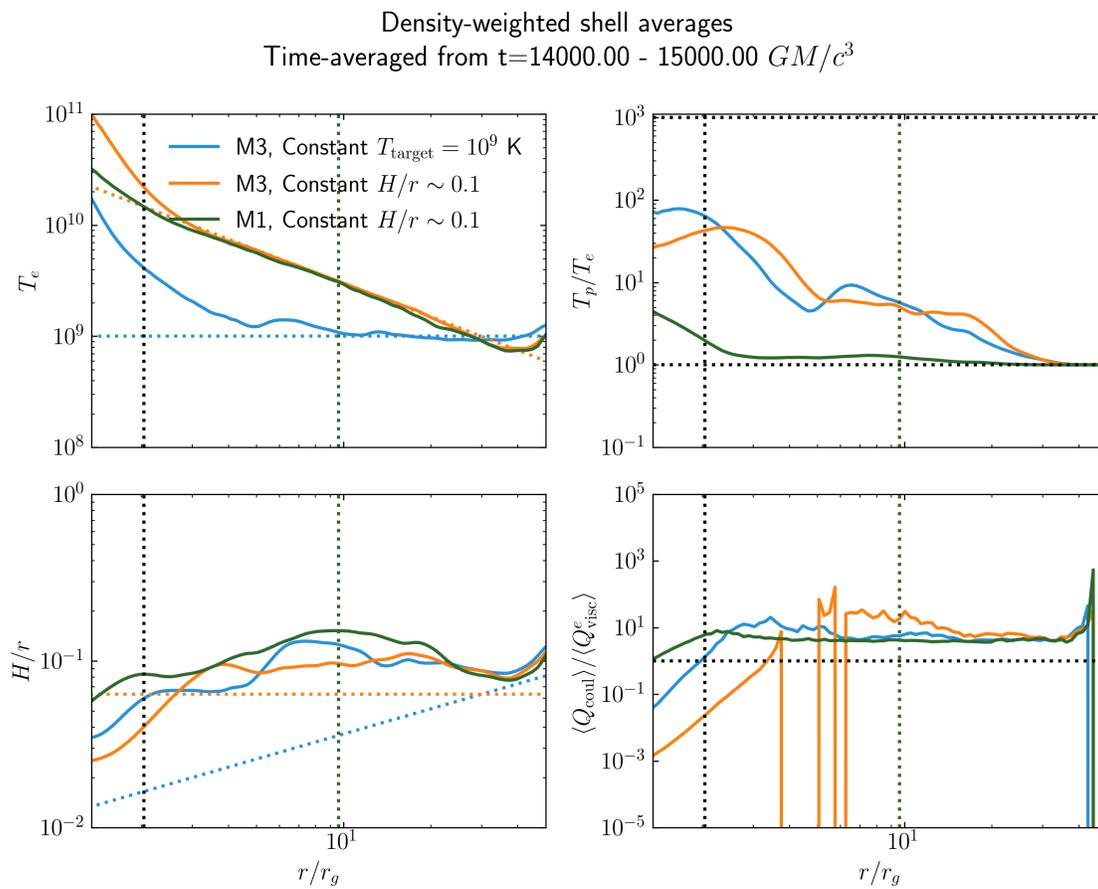


Figure 5.18: Density-weighted shell averages, time-averaged over the interval $1.4 \times 10^4 < tc/r_g < 1.5 \times 10^4$.

5.A Coulomb Heating Equilibrium Temperature

Because total energy is conserved, we can easily solve for the equilibrium temperature T_0 such that $T_e = T_i$ under the influence of Coulomb collisions. The temperature is related to the internal energy density as

$$T_k = A_k u_k, \quad (5.19)$$

where $A_p = (\gamma_p - 1)/(n_0 k_B)$ in SI units and $A_p = (\gamma_p - 1) * U_{\text{unit}}/(\rho * N_{\text{unit}} k_B)$ in code units. The electron energy density has a different coefficient A_e because it is found from the electron entropy:

$$u_e = \frac{\kappa_e \rho^{\gamma_e}}{\gamma_e - 1} \quad (5.20)$$

$$T_e = \frac{m_p c^2 (\gamma_e - 1)}{k_B \rho} u_e = A_e u_e. \quad (5.21)$$

Using Eq. 5.1, with $T_e = T_p = T_0$ we have

$$T_p = A_p (u_g - u_e) = A_p \left(u_g - \frac{T_e}{A_e} \right) \quad (5.22)$$

$$= T_0 = A_p \left(u_g - \frac{T_0}{A_e} \right). \quad (5.23)$$

Now solving for T_0 , we obtain the equilibrium temperature in terms of the internal energy density u_g :

$$T_0 = u_g \frac{A_e A_p}{A_e + A_p}. \quad (5.24)$$

This equation matches with Ref. [103] Eq. 13.

5.B Thin Disk Scale Height

To calculate the expected value and scaling of the scale height to radius ratio, we start with

$$H = \frac{c_s}{\Omega} \quad (5.25)$$

where the sound speed c_s is defined as $c_s^2 = \gamma P / \rho$ and we set the angular velocity to its Keplerian value $\Omega = \Omega_K = 1/(r^{1/5} + a)$. Reducing the sound speed more, we have

$$c_s^2 = \gamma_p \frac{P}{\rho} = \gamma_p \frac{n k_B T_p}{\rho} = \gamma_p m_p k_B T_p. \quad (5.26)$$

Plugging into Eq. 5.25, we have

$$\frac{H}{r} = \sqrt{\frac{k_B \gamma_p}{GM m_p}} T_p^{1/2} r_{\text{cm}}^{1/2} \quad (5.27)$$

$$= \sqrt{\frac{k_B \gamma_p}{m_p c^2}} T_p^{1/2} r_{\text{rg}}^{1/2} \quad (5.28)$$

where the first line is in SI units and the second is in code units.

For a black hole spin of $a = 0.9375$, we find that at the ISCO ($r = 2.0r_g$),

$$\frac{H}{r}(T = 10^9 \text{ K}) = 0.018 \quad (5.29)$$

$$\frac{H}{r}(T = 10^{10} \text{ K}) = 0.057 \quad (5.30)$$

$$(5.31)$$

5.C Implementing Isothermal Electrons

As a way to bracket the uncertainty in the viscous MHD heating, we want to implement isothermal electrons. The idea is that if we fix the electron temperature, we can see how viscous heating and Coulomb heating differs from the consistently-evolved electron temperature and obtain an upper bound for the cooling that occurs. There are multiple conceivable ways of implementing isothermal electrons: i) set the cooling time to be dt (i.e. instantaneous cooling), or ii) set the electrons to be isothermal at the target temperature. The first approach resulted in unphysical heating/cooling. That makes sense because the cooling function is known to be unstable when the cooling time is close to the time step value (e.g. Ref. [99, Sec. 3.3]). Therefore, we pursue the second option. Note that to actually set the electrons to be isothermal, we do not want to touch any of the fluid evolution, i.e. we only want to edit KEL, not KTOT. Also note that most of this work was conducted with the explicit Coulomb implementation. It may be beneficial in the future to examine isothermal electrons once the implicit Coulomb function implementation is fully tested.

To bracket the error due to viscous heating and our cooling function, there are two separate concepts that we test separately. The first concept is that of isothermal electrons, i.e. fixing the electron temperature at basically every step. The problem with that approach is that the output of Q_{coul} becomes essentially meaningless (if not just zero always). The second idea is related to attempting to fix the crash at high

accretion rates that is possibly due to Q_{coul} being large enough to trigger the T_p/T_e ceiling/floor – the Coulomb time u_e/Q_{coul} is on the order of the simulation time step. In this approach, we simply use the electron target temperature to calculate the Coulomb energy exchange rate Q_{coul} , but do not actually reset T_e at all. The benefit of this approach is that we can quantify the error — we know how much of a mistake we make by setting $T_e = T_{e,\text{target}}$ because we output the actual T_e . We can then calculate what Q_{coul} would have been if we had not made this assumption.

The isothermal electron implementation is flagged with the ISOTHERMALE tag, default False set in the `config.py` file. Setting T_e for Q_{coul} and nothing else is flagged as FIXEDTE. The idea is to just fix T_e in the Q_{coul} method, and then fix KEL after applying the radiation force to the electrons. The goal is to keep having a nonzero Q_{cool} and thus nonzero radiation pressure, but still having isothermal electrons. To enforce isothermal electrons, we make changes in the `electrons.c` file, flagged by the ISOTHERMALE tag. The very first change is in the `init electrons` function. The default initialization sets the electron temperature by assuming that the electron internal energy is a fixed fraction `fe10` of the total gas internal energy. Adjusting the electron temperature to have the target temperature does decrease the gas temperature (because $u_g = u_e + u_p$ is fixed and we're increasing u_e). In a test, the electron temperature increased from 3.38×10^8 K to 10^9 K and the proton temperature decreased from 6.7×10^{10} K to 6.6×10^9 K. However, in tests done later, the electron temperature could not be set to the target because it was hitting the T_p/T_e floor (electrons much hotter than protons).

5.C.1 Using Fixed Electron Temperature in Coulomb Exchange

As seen in Fig. 5.19, the fake Q_{coul} is fine for low \dot{m} , although it does have an error of almost 100% in the inner $2r_g$ (without fake Q_{coul} , $T_e = 3 \times 10^9$ K, with fake Q_{coul} , $T_e = 5 \times 10^9$ K). The density ρ is larger with the fake Q_{coul} because Q_{coul} is larger. Why is Q_{coul} larger? Well, if T_p/T_e is the same but T_e is 2 times larger, then $T_p - T_e$ and hence Q_{coul} is also two times larger:

$$T_{pb} - T_{eb} = T_{pb} \left(1 - \frac{T_{eb}}{T_{pb}}\right) = T_{pb} \left(1 - \frac{T_{ea}}{T_{pa}}\right) \quad (5.32)$$

$$= \frac{T_{pb}}{T_{pa}} (T_{ea} - T_{pa}). \quad (5.33)$$

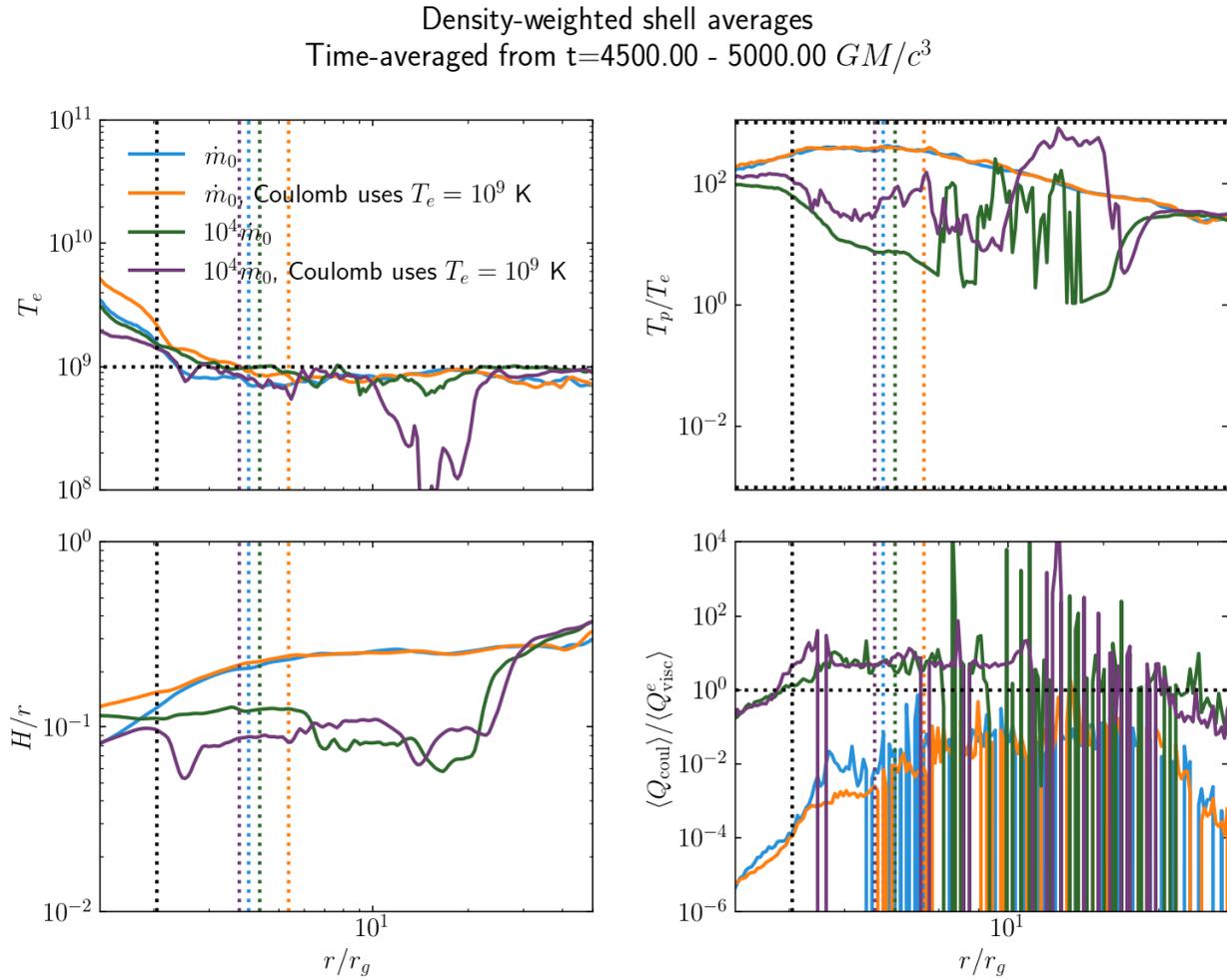


Figure 5.19: Density-weighted shell averages from 2D GRMHD torus tests with isothermal electrons, time-averaged over $4500 < tc/r_g < 5000$. This implementation sets $T_e = 10^9$ K only while calculating the Coulomb energy exchange rate. It does not fix the problems at high accretion rate, which are likely due to using an explicit evolution for the Coulomb heating (Ch. 5.4.3).

At the initial condition, the fake Q_{coul} results in a larger T_e than we would have otherwise ($T_e = 10^9$ instead of 10^8 K), which lowers T_p and makes $T_p - T_e$ smaller, i.e. Q_{coul} smaller. Then smaller Q_{coul} allows T_e and T_p to drift further apart. I actually think this implementation makes the problem worse, because ρ being higher than it should be results in higher Q_{coul} , which triggers the T_p/T_e ceiling/floor cycle faster. Indeed, the T_p/T_e ratio looks worse for the high \dot{m} fake Q_{coul} – it is basically at the ceiling at $r \sim 20r_g$.

5.C.2 Hard-coding $T_e = 10^9$ K

From Fig. 5.20, hardcoding the electron temperature whenever possible works fine for low \dot{m} : H/r looks very similar. T_p/T_e is much larger for the isothermal electrons, i.e. the protons are hotter. In Ch. 5.4.1, we see that for a given proton temperature, Q_{coul} is smaller for larger T_e . Since the initial condition changes to have hotter electrons, maybe that is decreasing Q_{coul} .

For larger accretion rate, the problem of T_p/T_e going wild seems to be solved. The purple T_p/T_e looks extremely similar to the small accretion rate isothermal case. Interestingly, H/r doesn't seem to change with accretion rate except within the ISCO, which may or may not be physical. Note that Q_{coul} is much larger, but the protons are still basically virial. The Coulomb quality factor looks good, so Q_{coul} is probably under control, i.e. it is circumventing the T_p/T_e ceiling/floor cycle. I had reset KEL after Q_{coul} , so Q_{coul} was super low and didn't back react on the gas at all, which could be the problem. Further tests will suss out what is wrong with this implementation.

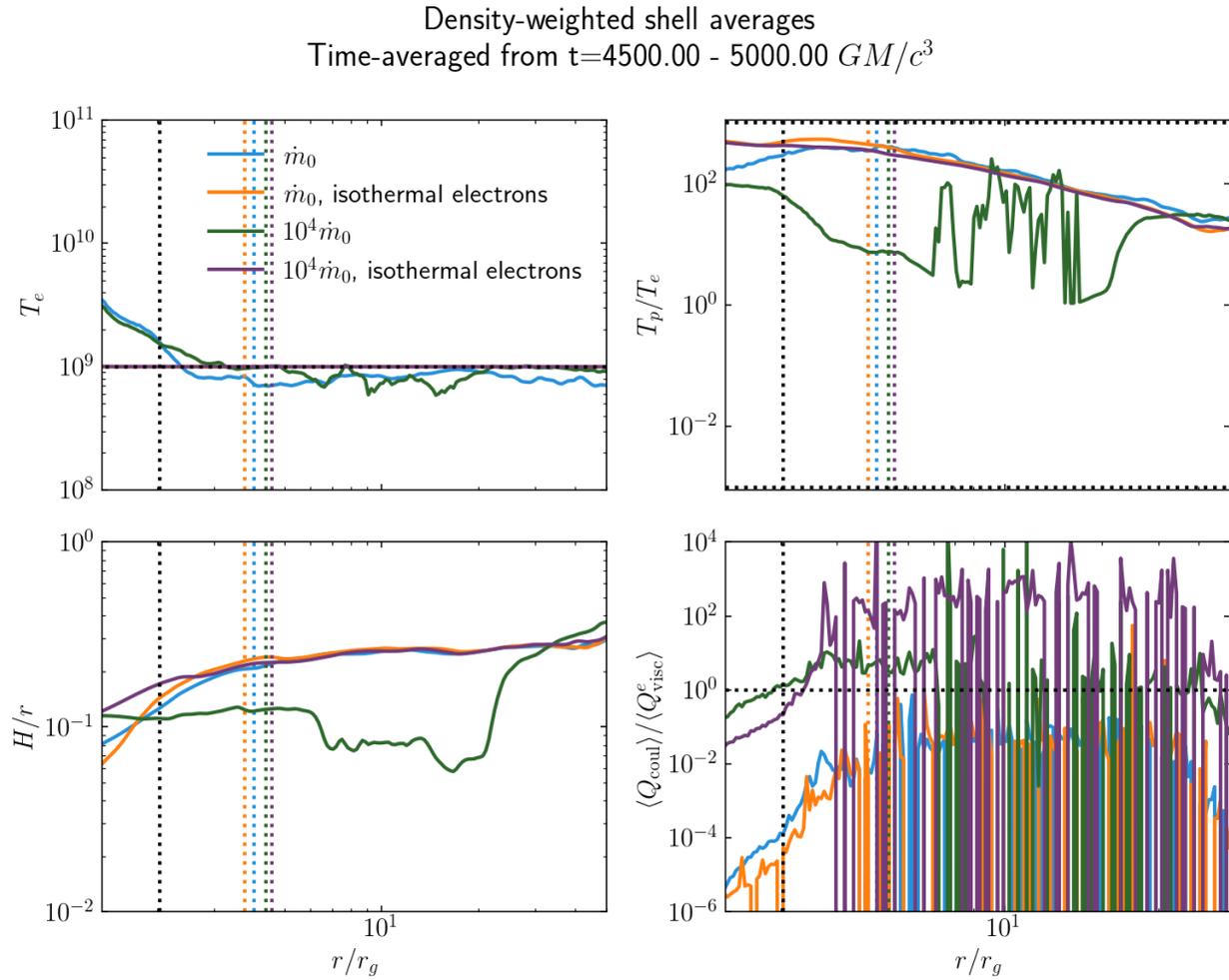


Figure 5.20: Density-weighted shell averages from 2D GRMHD torus tests, time-averaged over $4500 < tc/r_g < 5000$. This implementation sets $T_e = 10^9$ K only while calculating the Coulomb energy exchange rate. Because Q_{coul} is extremely large with seemingly no impact on the disk structure and H/r , something must be wrong with this implementation.

Chapter 6

Summary and Conclusions

In this thesis, I study two-temperature and kinetic plasma physics processes in the accretion disks and coronae around black holes. These processes can modify the underlying electron distribution and therefore affect the nonthermal emission observed from accretion disks and coronae around black holes. In principle, these processes can also affect the large-scale dynamics of accretion disks and their coronae, as discussed below. In the three main chapters of this thesis, I explore these processes on progressively larger length scales, moving from kinetic physics to MHD physics.

I begin by studying a small, local patch of plasma in the corona (Ch. 3). I ignore gradients in background density and magnetic field in favor of probing the kinetic physics of the collisionless coronal plasma with PIC simulations. To emulate how a corona sitting atop an accretion disk would receive energy from the disk, I drive the system with asymmetric energy. I examine how the turbulent cascade's formation is slowed by this asymmetry and demonstrate that particle acceleration also proceeds on slower timescales for this system. If the light-crossing time of the corona's finite extent is shorter than the development time of these high-energy particles, then the high-energy particles will not occur in the corona. In addition, I discover a momentum-coupling mechanism that could launch relativistic winds from a corona. Future work will determine whether this relativistic mechanism could affect the dynamics of a disk/corona system by, for example, removing a significant amount of angular momentum.

Next, I discuss the possibility of accelerating particles within the plunging region of a thin accretion disk (Ch. 4). I show that conditions are suitable for electrons to be accelerated to high-energies and to remain at high energies until they radiate away their energy, rather than being thermalized. In addition, I

use ray-tracing to show that power-law emission from within the plunging region comprises a fraction of the total luminosity that is consistent with observations. I make predictions for observational trends of the power-law fraction with black hole spin and the system inclination angle.

Finally, I simulate the two-temperature plasma in a full accretion disk (Ch. 5). I test the previous chapter's prediction that electrons and protons should decouple in the inner regions of an accretion disk when Coulomb collisions become inefficient. This decoupling could significantly affect the structure of an accretion disk, leading to the collapse of thick accretion flow or leading to a puffy disk at the inner parts of a thin accretion disk. By using full GRMHD simulations, I can probe the radial structure of this transition and how it depends on accretion rate. This work could lead to better understanding of the disk truncation model for BHB state transitions.

Moving forward, there are currently two main avenues for investigating collisionless effects on scales of an accretion disk. In the first, PIC simulations create prescriptions that GRMHD simulations can use to model a nonthermal population of electrons atop a thermal fluid. This approach is useful to gain a global understanding of how much small-scale physics should affect overall dynamics and observed radiation. However, this avenue is slightly tricky due to the number of important variables, including currently unknown parameters and transport of nonthermal particles. For example, my work in Ch. 3 showed that asymmetric energy injection can significantly impact the development of nonthermal electrons. So far, GRMHD simulations using PIC prescriptions have not included the effect of asymmetric energy injection, among other factors. Electron cooling in particular can drastically change the efficiency of particle acceleration, as can the presence of a guide field during reconnection. The other avenue is to use PIC for the entire accretion disk. This approach is extremely difficult computationally, and as such is still in the very early stages of development.

Upcoming X-ray telescopes will provide more opportunities for interpreting emission from the disk/corona system. Understanding the kinetic and two-temperature plasma physics of the disk/corona is crucial for creating accurate observational models. X-ray polarization measurements are often cited as a possible mechanism to distinguish between various coronal geometries, especially with the recent launch of IXPE. However, these models make assumptions about the disk/corona geometry in the first place, and

often assume a thermal electron distribution. In addition to IXPE, the soon-to-launch XRISM telescope will provide energy spectra with excellent spectral resolution. The hope is that XRISM will separate out spectroscopic lines, particularly the iron lines that are commonly used to measure black hole spin. However, the presence of these iron lines depends on an understanding of the ambient radiation distribution. The radiation in turn depends on the kinetic physics and dynamics in the accretion disk/corona. In addition, non-optically thin effects could completely change the interpretation of reverberation mapping data. Understanding where in the accretion disk different regimes apply is crucial to interpreting a variety of observational data. In the future, I hope to explore more ways to bridge the collisional and collisionless plasma physics regimes in the plasmas around black holes.

References

- [1] M. C. Weisskopf, B. Ramsey, S. O'Dell, A. Tennant, R. Elsner, P. Soffitta, R. Bellazzini, E. Costa, J. Kolodziejczak, V. Kaspi, F. Muleri, H. Marshall, G. Matt, and R. Romani, "The Imaging X-ray Polarimetry Explorer (IXPE)", [9905 \(2016\)](#).
- [2] XRISM Science Team, **Science with the X-ray Imaging and Spectroscopy Mission (XRISM)**, tech. rep., Publication Title: arXiv e-prints ADS Bibcode: 2020arXiv200304962X Type: article (Mar. 2020).
- [3] C. S. Reynolds, "Observational Constraints on Black Hole Spin", en, [Annual Review of Astronomy and Astrophysics, Volume 59, pp. 117-154 59 \(2021\)](#).
- [4] A. M. Hankla, V. Zhdankin, G. R. Werner, D. A. Uzdensky, and M. C. Begelman, "Kinetic simulations of imbalanced turbulence in a relativistic plasma: Net flow and particle acceleration", [Monthly Notices of the Royal Astronomical Society 509 \(2022\)](#).
- [5] A. M. Hankla, N. Scepi, and J. Dexter, "Non-thermal emission from the plunging region: a model for the high-energy tail of black hole X-ray binary soft states", [Monthly Notices of the Royal Astronomical Society 515 \(2022\)](#).
- [6] A. M. Hankla, N. Scepi, and J. Dexter, "Formation of a truncated disk due to Coulomb collisions",
- [7] A. M. Hankla, Y.-F. Jiang, and P. J. Armitage, "Local Simulations of Heating Torques on a Luminous Body in an Accretion Disk", [The Astrophysical Journal 902 \(2020\)](#).
- [8] T. Kroker, M. Großmann, K. Sengstock, M. Drescher, P. Wessels-Staarmann, and J. Simonet, "Ultrafast electron cooling in an expanding ultracold plasma", en, [Nature Communications 12 \(2021\)](#).

- [9] N. I. Shakura and R. A. Sunyaev, “Reprint of 1973A&A....24..337S. Black holes in binary systems. Observational appearance.”, **500** (1973).
- [10] A. B. Zylstra, O. A. Hurricane, D. A. Callahan, A. L. Kritcher, J. E. Ralph, H. F. Robey, J. S. Ross, C. V. Young, K. L. Baker, D. T. Casey, T. Döppner, L. Divol, M. Hohenberger, S. Le Pape, A. Pak, P. K. Patel, R. Tommasini, S. J. Ali, P. A. Amendt, L. J. Atherton, B. Bachmann, D. Bailey, L. R. Benedetti, L. Berzak Hopkins, R. Betti, S. D. Bhandarkar, J. Biener, R. M. Bionta, N. W. Birge, E. J. Bond, D. K. Bradley, T. Braun, T. M. Briggs, M. W. Bruhn, P. M. Celliers, B. Chang, T. Chapman, H. Chen, C. Choate, A. R. Christopherson, D. S. Clark, J. W. Crippen, E. L. Dewald, T. R. Dittrich, M. J. Edwards, W. A. Farmer, J. E. Field, D. Fittinghoff, J. Frenje, J. Gaffney, M. Gatu Johnson, S. H. Glenzer, G. P. Grim, S. Haan, K. D. Hahn, G. N. Hall, B. A. Hammel, J. Harte, E. Hartouni, J. E. Heebner, V. J. Hernandez, H. Herrmann, M. C. Herrmann, D. E. Hinkel, D. D. Ho, J. P. Holder, W. W. Hsing, H. Huang, K. D. Humbird, N. Izumi, L. C. Jarrott, J. Jeet, O. Jones, G. D. Kerbel, S. M. Kerr, S. F. Khan, J. Kilkenney, Y. Kim, H. Geppert Kleinrath, V. Geppert Kleinrath, C. Kong, J. M. Koning, J. J. Kroll, M. K. G. Kruse, B. Kustowski, O. L. Landen, S. Langer, D. Larson, N. C. Lemos, J. D. Lindl, T. Ma, M. J. MacDonald, B. J. MacGowan, A. J. Mackinnon, S. A. MacLaren, A. G. MacPhee, M. M. Marinak, D. A. Mariscal, E. V. Marley, L. Masse, K. Meaney, N. B. Meezan, P. A. Michel, M. Millot, J. L. Milovich, J. D. Moody, A. S. Moore, J. W. Morton, T. Murphy, K. Newman, J.-M. G. Di Nicola, A. Nikroo, R. Nora, M. V. Patel, L. J. Pelz, J. L. Peterson, Y. Ping, B. B. Pollock, M. Ratledge, N. G. Rice, H. Rinderknecht, M. Rosen, M. S. Rubery, J. D. Salmonson, J. Sater, S. Schiaffino, D. J. Schlossberg, M. B. Schneider, C. R. Schroeder, H. A. Scott, S. M. Sepke, K. Sequoia, M. W. Sherlock, S. Shin, V. A. Smalyuk, B. K. Spears, P. T. Springer, M. Stadermann, S. Stoupin, D. J. Strozzi, L. J. Suter, C. A. Thomas, R. P. J. Town, E. R. Tubman, C. Trosseille, P. L. Volegov, C. R. Weber, K. Widmann, C. Wild, C. H. Wilde, B. M. Van Wonterghem, D. T. Woods, B. N. Woodworth, M. Yamaguchi, S. T. Yang, and G. B. Zimmerman, “Burning plasma achieved in inertial fusion”, en, *Nature* **601** (2022).
- [11] M. W. Kunz, **Lecture Notes on Introduction to Plasma Astrophysics**, Nov. 2020.

- [12] D. Nicholson, **Introduction to Plasma Theory** (John Wiley and Sons, Inc., New York, 1983).
- [13] A. N. Kolmogorov, “The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers”, [Proceedings of the Royal Society of London Series A](#) **434** (1991).
- [14] P. Goldreich and S. Sridhar, “Toward a theory of interstellar turbulence. II. Strong alfvénic turbulence”, **438** (1995).
- [15] S. Boldyrev, “Spectrum of magnetohydrodynamic turbulence”, **96** (2006).
- [16] C. Dong, L. Wang, Y.-M. Huang, L. Comisso, T. A. Sandstrom, and A. Bhattacharjee, “Reconnection-driven energy cascade in magnetohydrodynamic turbulence”, [Science Advances](#) **8** (2022).
- [17] A. A. Schekochihin, “MHD turbulence: A biased review”, arXiv e-prints (2020).
- [18] A. Lazarian, G. L. Eyink, A. Jafari, G. Kowal, H. Li, S. Xu, and E. T. Vishniac, “3D turbulent reconnection: Theory, tests, and astrophysical implications”, [Physics of Plasmas](#) **27** (2020).
- [19] J. C. Perez and S. Boldyrev, “Numerical simulations of imbalanced strong magnetohydrodynamic turbulence”, **710** (2010).
- [20] J. C. Perez and S. Boldyrev, “Strong magnetohydrodynamic turbulence with cross helicity”, [Physics of Plasmas](#) **17** (2010).
- [21] Y. Lithwick, P. Goldreich, and S. Sridhar, “Imbalanced Strong MHD Turbulence”, [ApJ](#) **655** (2007).
- [22] A. Beresnyak and A. Lazarian, “MHD turbulence, turbulent dynamo and applications”, in **Magnetic fields in diffuse media**, Vol. 407, edited by A. Lazarian, E. M. de Gouveia Dal Pino, and C. Melioli, Astrophysics and space science library, tex.adsnote: Provided by the SAO/NASA Astrophysics Data System tex.adsurl: <https://ui.adsabs.harvard.edu/abs/2015ASSL..407..163B> (2015), p. 163.
- [23] B. D. G. Chandran, “Strong anisotropic MHD turbulence with cross helicity”, **685** (2008).
- [24] A. Beresnyak and A. Lazarian, “Strong imbalanced turbulence”, **682** (2008).
- [25] A. Beresnyak and A. Lazarian, “Structure of stationary strong imbalanced turbulence”, **702** (2009).
- [26] A. Beresnyak and A. Lazarian, “Scaling laws and diffuse locality of balanced and imbalanced magnetohydrodynamic turbulence”, **722** (2010).

- [27] J. Cho and A. Lazarian, “Imbalanced Relativistic Force-free Magnetohydrodynamic Turbulence”, *\apj* **780** (2014).
- [28] J. C. Perez, J. Mason, S. Boldyrev, and F. Cattaneo, “On the energy spectrum of strong magnetohydrodynamic turbulence”, *Physical Review X* **2** (2012).
- [29] J. C. Perez and S. Boldyrev, “Role of cross-helicity in magnetohydrodynamic turbulence”, **102** (2009).
- [30] J. Mason, J. C. Perez, S. Boldyrev, and F. Cattaneo, “Numerical simulations of strong incompressible magnetohydrodynamic turbulence”, *Physics of Plasmas* **19** (2012).
- [31] S. A. Markovskii and B. J. Vasquez, “Magnetic helicity in the dissipation range of strong imbalanced turbulence”, **768** (2013).
- [32] H. Kim and J. Cho, “Inverse cascade in imbalanced electron magnetohydrodynamic turbulence”, **801** (2015).
- [33] Y. Voitenko and J. De Keyser, “MHD-Kinetic transition in imbalanced alfvénic turbulence”, **832** (2016).
- [34] G. Miloshevich, T. Passot, and P. L. Sulem, “Modeling imbalanced collisionless alfvén wave turbulence with nonlinear diffusion equations”, **888** (2020).
- [35] T. Passot and P. L. Sulem, “Imbalanced kinetic Alfvén wave turbulence: from weak turbulence theory to nonlinear diffusion models for the strong regime”, *Journal of Plasma Physics* **85** (2019).
- [36] N. J. Fox, M. C. Velli, S. D. Bale, R. Decker, A. Driesman, R. A. Howard, J. C. Kasper, J. Kinnison, M. Kusterer, D. Lario, M. K. Lockwood, D. J. McComas, N. E. Raouafi, and A. Szabo, “The solar probe plus mission: Humanity’s first visit to our star”, **204** (2016).
- [37] S. A. Fuselier, W. S. Lewis, C. Schiff, R. Ergun, J. L. Burch, S. M. Petrinec, and K. J. Trattner, “Magnetospheric multiscale science mission profile and operations”, **199** (2016).
- [38] J. W. Belcher and J. Davis Leverett, “Large-amplitude Alfvén waves in the interplanetary medium, 2”, **76** (1971).

- [39] R. Bruno and V. Carbone, “The solar wind as a turbulence laboratory”, [Living Reviews in Solar Physics](#) **10** (2013).
- [40] W. H. Matthaeus, G. P. Zank, S. Oughton, D. J. Mullan, and P. Dmitruk, “Coronal Heating by Magnetohydrodynamic Turbulence Driven by Reflected Low-Frequency Waves”, [ApJ](#) **523** (1999).
- [41] Y. M. Voitenko, “Excitation of Kinetic Alfvén Waves in a Flaring Loop”, [SolPhys](#) **182** (1998).
- [42] Y. Voitenko and M. Goossens, “Competition of damping mechanisms for the phase-mixed Alfvén waves in the solar corona”, [ApJ](#) **357** (2000).
- [43] C. H. K. Chen, S. D. Bale, C. S. Salem, and B. A. Maruca, “Residual energy spectrum of solar wind turbulence”, [ApJ](#) **770** (2013).
- [44] D. Vech, J. C. Kasper, K. G. Klein, J. Huang, M. L. Stevens, C. H. K. Chen, A. W. Case, K. Korreck, S. D. Bale, T. A. Bowen, P. L. Whittlesey, R. Livi, D. E. Larson, D. Malaspina, M. Pulupa, J. Bonnell, P. Harvey, K. Goetz, T. Dudok de Wit, and R. MacDowall, “Kinetic-scale spectral features of cross helicity and residual energy in the inner heliosphere”, [ApJ](#) **246** (2020).
- [45] C. H. K. Chen, “Recent progress in astrophysical plasma turbulence from solar wind observations”, [Journal of Plasma Physics](#) **82** (2016).
- [46] J. Frank, A. King, and D. J. Raine, **Accretion Power in Astrophysics: Third Edition** (Cambridge University Press, 2002).
- [47] C. Bambi, L. W. Brenneman, T. Dauser, J. A. García, V. Grinberg, A. Ingram, J. Jiang, H. Liu, A. M. Lohfink, A. Marinucci, G. Mastroserio, R. Middei, S. Nampalliwar, A. Niedźwiecki, J. F. Steiner, A. Tripathi, and A. A. Zdziarski, “Towards Precision Measurements of Accreting Black Holes Using X-Ray Reflection Spectroscopy”, en [Space Science Reviews](#) **217** (2021).
- [48] R. A. Remillard and J. E. McClintock, “X-ray properties of black-hole binaries”, [ApJ](#) **44** (2006).
- [49] I. D. Novikov and K. S. Thorne, “Astrophysics of black holes.”, in *Black Holes (Les Astres Occlus)* (Jan. 1973), pp. 343–450.

- [50] E. Agol and J. H. Krolik, “Magnetic stress at the marginally stable orbit: Altered disk structure, radiation, and black hole spin evolution”, **528** (2000).
- [51] C. F. Gammie, “Efficiency of Magnetized Thin Accretion Disks in the Kerr Metric”, *\apjl* **522** (1999).
- [52] S. L. Shapiro, A. P. Lightman, and D. M. Eardley, “A two-temperature accretion disk model for Cygnus X-1: structure and spectrum.”, *The Astrophysical Journal* **204** (1976).
- [53] S. Ichimaru, “Bimodal behavior of accretion disks: theory and application to Cygnus X-1 transitions.”, *The Astrophysical Journal* **214** (1977).
- [54] M. J. Rees, M. C. Begelman, R. D. Blandford, and E. S. Phinney, “Ion-supported tori and the origin of radio jets”, **295** (1982).
- [55] M. A. Abramowicz, B. Czerny, J. P. Lasota, and E. Szuszkiewicz, “Slim Accretion Disks”, *The Astrophysical Journal* **332** (1988).
- [56] R. Narayan and I. Yi, “Advection-dominated Accretion: A Self-similar Solution”, *\apjl* **428** (1994).
- [57] G. B. Rybicki and A. P. Lightman, **Radiative processes in astrophysics**, tex.adsnote: Provided by the SAO/NASA Astrophysics Data System tex.adsurl: <https://ui.adsabs.harvard.edu/abs/1986rpa..book.....R> (John Wiley & Sons, Inc., 1979).
- [58] J. Dexter, N. Scepi, and M. C. Begelman, “Radiation GRMHD Simulations of the Hard State of Black Hole X-Ray Binaries and the Collapse of a Hot Accretion Flow”, *The Astrophysical Journal* **919** (2021).
- [59] R. Mahadevan and E. Quataert, “Are Particles in Advection-dominated Accretion Flows Thermal?”, *The Astrophysical Journal* **490** (1997).
- [60] S. A. Balbus and J. F. Hawley, “A Powerful Local Shear Instability in Weakly Magnetized Disks. I. Linear Analysis”, *\apj* **376** (1991).
- [61] S. A. Balbus and J. F. Hawley, “Instability, turbulence, and enhanced transport in accretion disks”, *Reviews of Modern Physics* **70** (1998).

- [62] R. D. Blandford and D. G. Payne, “Hydromagnetic flows from accretion disks and the production of radio jets.”, *Monthly Notices of the Royal Astronomical Society* **199** (1982).
- [63] R. F. Penna, J. C. McKinney, R. Narayan, A. Tchekhovskoy, R. Shafee, and J. E. McClintock, “Simulations of magnetized discs around black holes: effects of black hole spin, disc thickness and magnetic field geometry”, *Monthly Notices of the Royal Astronomical Society* **408** (2010).
- [64] S. C. Noble, J. H. Krolik, and J. F. Hawley, “Dependence of Inner Accretion Disk Stress on Parameters: The Schwarzschild Case”, *The Astrophysical Journal* **711** (2010).
- [65] R. F. Penna, A. Sądowski, and J. C. McKinney, “Thin-disc theory with a non-zero-torque boundary condition and comparisons with simulations”, *Monthly Notices of the Royal Astronomical Society* **420** (2012).
- [66] G. G. Howes, “A prescription for the turbulent heating of astrophysical plasmas”, *Monthly Notices of the Royal Astronomical Society* **409** (2010).
- [67] M. E. Rowan, L. Sironi, and R. Narayan, “Electron and Proton Heating in Transrelativistic Magnetic Reconnection”, *The Astrophysical Journal* **850** (2017).
- [68] G. R. Werner, D. A. Uzdensky, M. C. Begelman, B. Cerutti, and K. Nalewajko, “Non-thermal particle acceleration in collisionless relativistic electron-proton reconnection”, *Monthly Notices of the Royal Astronomical Society* **473** (2018).
- [69] A. Chael, R. Narayan, and M. D. Johnson, “Two-temperature, Magnetically Arrested Disc simulations of the jet from the supermassive black hole in M87”, *Monthly Notices of the Royal Astronomical Society* **486** (2019).
- [70] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, J. C. Algaba, R. Anantua, K. Asada, R. Azulay, U. Bach, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett, M. Bauböck, B. A. Benson, D. Bintley, L. Blackburn, R. Blundell, K. L. Bouman, G. C. Bower, H. Boyce, M. Bremer, C. D. Brinkerink, R. Brissenden, S. Britzen, A. E. Broderick, D. Brogiere, T. Bronzwaer, S. Bustamante, D.-Y. Byun, J. E. Carlstrom, C. Ceccobello, A. Chael, C.-k. Chan, K. Chatterjee,

S. Chatterjee, M.-T. Chen, Y. Chen, X. Cheng, I. Cho, P. Christian, N. S. Conroy, J. E. Conway, J. M. Cordes, T. M. Crawford, G. B. Crew, A. Cruz-Osorio, Y. Cui, J. Davelaar, M. De Laurentis, R. Deane, J. Dempsey, G. Desvignes, J. Dexter, V. Dhruv, S. S. Doeleman, S. Dougal, S. A. Dzib, R. P. Eatough, R. Emami, H. Falcke, J. Farah, V. L. Fish, E. Fomalont, H. A. Ford, R. Fraga-Encinas, W. T. Freeman, P. Friberg, C. M. Fromm, A. Fuentes, P. Galison, C. F. Gammie, R. García, O. Gentaz, B. Georgiev, C. Goddi, R. Gold, A. I. Gómez-Ruiz, J. L. Gómez, M. Gu, M. Gurwell, K. Hada, D. Haggard, K. Haworth, M. H. Hecht, R. Hesper, D. Heumann, L. C. Ho, P. Ho, M. Honma, C.-W. L. Huang, L. Huang, D. H. Hughes, S. Ikeda, C. M. Violette Impellizzeri, M. Inoue, S. Issaoun, D. J. James, B. T. Jannuzi, M. Janssen, B. Jeter, W. Jiang, A. Jiménez-Rosales, M. D. Johnson, S. Jorstad, A. V. Joshi, T. Jung, M. Karami, R. Karuppusamy, T. Kawashima, G. K. Keating, M. Kettenis, D.-J. Kim, J.-Y. Kim, J. Kim, J. Kim, M. Kino, J. Y. Koay, P. Kocherlakota, Y. Kofuji, P. M. Koch, S. Koyama, C. Kramer, M. Kramer, T. P. Krichbaum, C.-Y. Kuo, N. La Bella, T. R. Lauer, D. Lee, S.-S. Lee, P. K. Leung, A. Levis, Z. Li, R. Lico, G. Lindahl, M. Lindqvist, M. Lisakov, J. Liu, K. Liu, E. Liuzzo, W.-P. Lo, A. P. Lobanov, L. Loinard, C. J. Lonsdale, R.-S. Lu, J. Mao, N. Marchili, S. Markoff, D. P. Marrone, A. P. Marscher, I. Martí-Vidal, S. Matsushita, L. D. Matthews, L. Medeiros, K. M. Menten, D. Michalik, I. Mizuno, Y. Mizuno, J. M. Moran, K. Moriyama, M. Moscibrodzka, C. Müller, A. Mus, G. Musoke, I. Myserlis, A. Nadolski, H. Nagai, N. M. Nagar, M. Nakamura, R. Narayan, G. Narayanan, I. Natarajan, A. Nathanail, S. Navarro Fuentes, J. Neilsen, R. Neri, C. Ni, A. Noutsos, M. A. Nowak, J. Oh, H. Okino, H. Olivares, G. N. Ortiz-León, T. Oyama, F. Özel, D. C. M. Palumbo, G. Filippos Paraschos, J. Park, H. Parsons, N. Patel, U.-L. Pen, D. W. Pesce, V. Piétu, R. Plambeck, A. PopStefanija, O. Porth, F. M. Pötzl, B. Prather, J. A. Preciado-López, D. Psaltis, H.-Y. Pu, V. Ramakrishnan, R. Rao, M. G. Rawlings, A. W. Raymond, L. Rezzolla, A. Ricarte, B. Ripperda, F. Roelofs, A. Rogers, E. Ros, C. Romero-Cañizales, A. Roshanineshat, H. Rottmann, A. L. Roy, I. Ruiz, C. Ruszczyk, K. L. J. Rygl, S. Sánchez, D. Sánchez-Argüelles, M. Sánchez-Portal, M. Sasada, K. Satopathy, T. Savolainen, F. P. Schloerb, J. Schonfeld, K.-F. Schuster, L. Shao, Z. Shen, D. Small, B. W. Sohn, J. SooHoo, K. Souccar, H. Sun, F. Tazaki, A. J. Tetarenko, P. Tiede, R. P. J. Tilanus, M. Titus, P. Torne, E. Traianou, T. Trent, S. Trippe, M. Turk, I. van Bemmell, H. J. van Langevelde, D. R.

- van Rossum, J. Vos, J. Wagner, D. Ward-Thompson, J. Wardle, J. Weintroub, N. Wex, R. Wharton, M. Wielgus, K. Wiik, G. Witzel, M. F. Wondrak, G. N. Wong, Q. Wu, P. Yamaguchi, D. Yoon, A. Young, K. Young, Z. Younsi, F. Yuan, Y.-F. Yuan, J. A. Zensus, S. Zhang, G.-Y. Zhao, S.-S. Zhao, T. L. Chan, R. Qiu, S. Ressler, and C. White, “First Sagittarius A* Event Horizon Telescope Results. V. Testing Astrophysical Models of the Galactic Center Black Hole”, *The Astrophysical Journal* **930** (2022).
- [71] S. Markoff, M. A. Nowak, and J. Wilms, “Going with the Flow: Can the Base of Jets Subsume the Role of Compact Accretion Disk Coronae?”, *Apj* **635** (2005).
- [72] E. Kara, J. F. Steiner, A. C. Fabian, E. M. Cackett, P. Uttley, R. A. Remillard, K. C. Gendreau, Z. Arzoumanian, D. Altamirano, S. Eikenberry, T. Enoto, J. Homan, J. Nielsen, and A. L. Stevens, “The corona contracts in a black-hole transient”, *en*, *Nature* **565** (2019).
- [73] A. A. Galeev, R. Rosner, and G. S. Vaiana, “Structured coronae of accretion disks.”, *The Astrophysical Journal* **229** (1979).
- [74] F. Haardt and L. Maraschi, “A Two-Phase Model for the X-Ray Emission from Seyfert Galaxies”, *The Astrophysical Journal* **380** (1991).
- [75] A. M. Beloborodov, “Plasma ejection from magnetic flares and the X-ray spectrum of cygnus X-1”, **510** (1999).
- [76] A. A. Esin, J. E. McClintock, and R. Narayan, “Advection-Dominated Accretion and the Spectral States of Black Hole X-Ray Binaries: Application to Nova Muscae 1991”, *The Astrophysical Journal* **489** (1997).
- [77] A. C. Fabian, D. J. Buisson, P. Kosec, C. S. Reynolds, D. R. Wilkins, J. A. Tomsick, D. J. Walton, P. Gandhi, D. Altamirano, Z. Arzoumanian, E. M. Cackett, S. Dyda, J. A. Garcia, K. C. Gendreau, B. W. Grefenstette, J. Homan, E. Kara, R. M. Ludlam, J. M. Miller, and J. F. Steiner, “The soft state of the black hole transient source MAXI J1820+070: emission from the edge of the plunge region?”, *Monthly Notices of the Royal Astronomical Society* **493** (2020).

- [78] C. Done, M. Gierliński, and A. Kubota, “Modelling the behaviour of accretion flows in X-ray binaries. Everything you always wanted to know about accretion but were afraid to ask”, *Astronomy and Astrophysics Review* **15** (2007).
- [79] J. D. Hogg and C. S. Reynolds, “The dynamics of truncated black hole accretion disks. II. Magneto-hydrodynamic case”, **854** (2018).
- [80] M. Liska, G. Musoke, A. Tchekhovskoy, and O. Porth, “Formation of Magnetically Truncated Accretion Disks in 3D Radiation-Transport Two-Temperature GRMHD Simulations”, arXiv e-prints (2022).
- [81] R. Nemmen, A. Vemado, I. Almeida, J. Garcia, and P. Motta, **Emergence of hot corona and truncated disk in simulations of accreting stellar mass black holes**, arXiv:2305.11429 [astro-ph], May 2023.
- [82] B. F. Liu, R. E. Taam, E. Meyer-Hofmeister, and F. Meyer, “The Existence of Inner Cool Disks in the Low/Hard State of Accreting Black Holes”, *The Astrophysical Journal* **671** (2007).
- [83] S. R. Cranmer, A. A. van Ballegoijen, and R. J. Edgar, “Self-consistent coronal heating and solar wind acceleration from anisotropic magnetohydrodynamic turbulence”, **171** (2007).
- [84] B. D. G. Chandran, F. Foucart, and A. Tchekhovskoy, “Heating of accretion-disk coronae and jets by general relativistic magnetohydrodynamic turbulence”, *Journal of Plasma Physics* **84** (2018).
- [85] G. R. Werner and D. A. Uzdensky, “Nonthermal particle acceleration in 3D relativistic magnetic reconnection in pair plasma”, **843** (2017).
- [86] V. Zhdankin, G. R. Werner, D. A. Uzdensky, and M. C. Begelman, “Kinetic turbulence in relativistic plasma: From thermal bath to nonthermal continuum”, **118** (2017).
- [87] P. S. Coppi, “The Physics of Hybrid Thermal/Non-Thermal Plasmas”, **161** (1999).
- [88] A. A. Chael, R. Narayan, and A. Sądowski, “Evolving non-thermal electrons in simulations of black hole accretion”, en, *Monthly Notices of the Royal Astronomical Society* **470** (2017).

- [89] V. Zhdankin, D. A. Uzdensky, G. R. Werner, and M. C. Begelman, “Electron and Ion Energization in Relativistic Plasma Turbulence”, *ApJ* **122** (2019).
- [90] G. R. Werner and D. A. Uzdensky, “Reconnection and particle acceleration in three-dimensional current sheet evolution in moderately magnetized astrophysical pair plasma”, *Journal of Plasma Physics* **87** (2021).
- [91] Y. Kawazura, A. A. Schekochihin, M. Barnes, J. M. TenBarge, Y. Tong, K. G. Klein, and W. Dorland, “Ion versus Electron Heating in Compressively Driven Astrophysical Gyrokinetic Turbulence”, *Physical Review X* **10** (2020).
- [92] V. Zhdankin, “Particle Energization in Relativistic Plasma Turbulence: Solenoidal versus Compressive Driving”, *The Astrophysical Journal* **922** (2021).
- [93] R. Basak, A. A. Zdziarski, M. Parker, and N. Islam, “Analysis of NuSTAR and Suzaku observations of Cyg X-1 in the hard state: evidence for a truncated disc geometry”, *Monthly Notices of the Royal Astronomical Society* **472** (2017).
- [94] C. S. Reynolds and A. C. Fabian, “Broad Iron- $K\alpha$ Emission Lines as a Diagnostic of Black Hole Spin”, *The Astrophysical Journal* **675** (2008).
- [95] D. R. Wilkins, L. C. Gallo, E. Costantini, W. N. Brandt, and R. D. Blandford, “Acceleration and cooling of the corona during X-ray flares from the Seyfert galaxy I Zw 1”, *Monthly Notices of the Royal Astronomical Society* **512** (2022).
- [96] E. M. Cackett, M. C. Bentz, and E. Kara, “Reverberation mapping of active galactic nuclei: From X-ray corona to dusty torus”, *en, iScience* **24** (2021).
- [97] G. Salvesen, “An Electron-scattering Time Delay in Black Hole Accretion Disks”, *The Astrophysical Journal* **940** (2022).
- [98] S. C. Noble, C. F. Gammie, J. C. McKinney, and L. Del Zanna, “Primitive variable solvers for conservative general relativistic magnetohydrodynamics”, **641** (2006).

- [99] B. R. Ryan, J. C. Dolence, and C. F. Gammie, “Bhlight: General Relativistic Radiation Magnetohydrodynamics with Monte Carlo Transport”, *The Astrophysical Journal* **807** (2015).
- [100] B. R. Ryan, S. M. Ressler, J. C. Dolence, A. Tchekhovskoy, C. Gammie, and E. Quataert, “The Radiative Efficiency and Spectra of Slowly Accreting Black Holes from Two-temperature GRRMHD Simulations”, *The Astrophysical Journal* **844** (2017).
- [101] A. Sądowski, R. Narayan, A. Tchekhovskoy, and Y. Zhu, “Semi-implicit scheme for treating radiation under M1 closure in general relativistic conservative fluid dynamics codes”, *Monthly Notices of the Royal Astronomical Society* **429** (2013).
- [102] S. Stepney and P. W. Guilbert, “Numerical fits to important rates in high temperature astrophysical plasmas.”, **204** (1983).
- [103] S. M. Ressler, A. Tchekhovskoy, E. Quataert, M. Chandra, and C. F. Gammie, “Electron thermodynamics in GRMHD simulations of low-luminosity black hole accretion”, *Monthly Notices of the Royal Astronomical Society* **454** (2015).
- [104] A. Sądowski, M. Wielgus, R. Narayan, D. Abarca, J. C. McKinney, and A. Chael, “Radiative, two-temperature simulations of low-luminosity black hole accretion flows in general relativity”, *Monthly Notices of the Royal Astronomical Society* **466** (2017).
- [105] K. Parfrey, A. Philippov, and B. Cerutti, “First-Principles Plasma Simulations of Black-Hole Jet Launching”, *Physical Review Letters* **122** (2019).
- [106] B. Crinquand, B. Cerutti, A. Philippov, K. Parfrey, and G. Dubus, “Multidimensional Simulations of Ergospheric Pair Discharges around Black Holes”, *Physical Review Letters* **124** (2020).
- [107] A. Galishnikova, A. Philippov, E. Quataert, F. Bacchini, K. Parfrey, and B. Ripperda, “Collisionless Accretion onto Black Holes: Dynamics and Flares”, *Physical Review Letters* **130** (2023).
- [108] C. Birdsall and A. Langdon, **Plasma Physics Via Computer Simulation** (CRC Press, Boca Raton, 2005).

- [109] B. Cerutti, G. R. Werner, D. A. Uzdensky, and M. C. Begelman, “Simulations of particle acceleration beyond the classical synchrotron burnoff limit in magnetic reconnection: An explanation of the crab flares”, **770** (2013).
- [110] G. R. Werner, A. A. Philippov, and D. A. Uzdensky, “Particle acceleration in relativistic magnetic reconnection with strong inverse-Compton cooling in pair plasmas”, *Monthly Notices of the Royal Astronomical Society* **482** (2019).
- [111] D. Grosej, H. Hakobyan, A. M. Beloborodov, L. Sironi, and A. Philippov, **Ab Initio Turbulent Comptonization in Magnetized Coronae of Accreting Black Holes**, tech. rep., Publication Title: arXiv e-prints ADS Bibcode: 2023arXiv230111327G Type: article (Jan. 2023).
- [112] W. B. Sparks, J. A. Biretta, and F. Macchetto, “The jet of M87 at tenth-arcsecond resolution: Optical, ultraviolet, and radio observations”, **473** (1996).
- [113] J. J. Hester, “The crab nebula : an astrophysical chimera.”, **46** (2008).
- [114] V. Zhdankin, D. A. Uzdensky, G. R. Werner, and M. C. Begelman, “System-size convergence of nonthermal particle acceleration in relativistic plasma turbulence”, **867** (2018).
- [115] L. Comisso and L. Sironi, “Particle acceleration in relativistic plasma turbulence”, **121** (2018).
- [116] L. Comisso and L. Sironi, “The interplay of magnetically dominated turbulence and magnetic reconnection in producing nonthermal particles”, **886** (2019).
- [117] A. A. Abdo, M. Ackermann, M. Ajello, W. B. Atwood, M. Axelsson, L. Baldini, J. Ballet, G. Barbiellini, D. Bastieri, M. Battelino, B. M. Baughman, K. Bechtol, R. Bellazzini, B. Berenji, R. D. Blandford, E. D. Bloom, G. Bogaert, E. Bonamente, A. W. Borgland, J. Bregeon, A. Brez, M. Brigida, P. Bruel, T. H. Burnett, G. A. Caliandro, R. A. Cameron, P. A. Caraveo, P. Carlson, J. M. Casandjian, C. Cecchi, E. Charles, A. Chekhtman, C. C. Cheung, J. Chiang, S. Ciprini, R. Claus, J. Cohen-Tanugi, L. R. Cominsky, J. Conrad, S. Cutini, C. D. Dermer, A. de Angelis, F. de Palma, S. W. Digel, G. di Bernardo, E. Do Couto E Silva, P. S. Drell, R. Dubois, D. Dumora, Y. Edmonds, C. Farnier, C. Favuzzi, W. B. Focke, M. Frailis, Y. Fukazawa, S. Funk, P. Fusco, D. Gaggero, F.

Gargano, D. Gasparrini, N. Gehrels, S. Germani, B. Giebels, N. Giglietto, F. Giordano, T. Glanzman, G. Godfrey, D. Grasso, I. A. Grenier, M. -. Grondin, J. E. Grove, L. Guillemot, S. Guiriec, Y. Hanabata, A. K. Harding, R. C. Hartman, M. Hayashida, E. Hays, R. E. Hughes, G. Jóhannesson, A. S. Johnson, R. P. Johnson, W. N. Johnson, T. Kamae, H. Katagiri, J. Kataoka, N. Kawai, M. Kerr, J. Knödlseider, D. Kocevski, F. Kuehn, M. Kuss, J. Lande, L. Latronico, M. Lemoine-Goumard, F. Longo, F. Loparco, B. Lott, M. N. Lovellette, P. Lubrano, G. M. Madejski, A. Makeev, M. M. Massai, M. N. Mazziotta, W. McConville, J. E. McEnery, C. Meurer, P. F. Michelson, W. Mitthumsiri, T. Mizuno, A. A. Moiseev, C. Monte, M. E. Monzani, E. Moretti, A. Morselli, I. V. Moskalenko, S. Murgia, P. L. Nolan, J. P. Norris, E. Nuss, T. Ohsugi, N. Omodei, E. Orlando, J. F. Ormes, M. Ozaki, D. Paneque, J. H. Panetta, D. Parent, V. Pelassa, M. Pepe, M. Pesce-Rollins, F. Piron, M. Pohl, T. A. Porter, S. Profumo, S. Rainò, R. Rando, M. Razzano, A. Reimer, O. Reimer, T. Reposeur, S. Ritz, L. S. Rochester, A. Y. Rodriguez, R. W. Romani, M. Roth, F. Ryde, H. F. -. Sadrozinski, D. Sanchez, A. Sander, P. M. Saz Parkinson, J. D. Scargle, T. L. Schalk, A. Sellerholm, C. Sgrò, D. A. Smith, P. D. Smith, G. Spandre, P. Spinelli, J. -. Starck, T. E. Stephens, M. S. Strickman, A. W. Strong, D. J. Suson, H. Tajima, H. Takahashi, T. Takahashi, T. Tanaka, J. B. Thayer, J. G. Thayer, D. J. Thompson, L. Tibaldo, O. Tibolla, D. F. Torres, G. Tosti, A. Tramacere, Y. Uchiyama, T. L. Usher, A. van Etten, V. Vasileiou, N. Vilchez, V. Vitale, A. P. Waite, E. Wallace, P. Wang, B. L. Winer, K. S. Wood, T. Ylinen, and M. Ziegler, “Measurement of the cosmic ray e^+e^- spectrum from 20GeV to 1TeV with the fermi large area telescope”, **102** (2009).

- [118] J. Aleksić, S. Ansoldi, L. A. Antonelli, P. Antoranz, A. Babic, P. Bangale, J. A. Barrio, J. Becerra González, W. Bednarek, E. Bernardini, B. Biasuzzi, A. Biland, O. Blanch, S. Bonnefoy, G. Bonoli, F. Borracci, T. Bretz, E. Carmona, A. Carosi, P. Colin, E. Colombo, J. L. Contreras, J. Cortina, S. Covino, P. Da Vela, F. Dazzi, A. De Angelis, G. De Caneva, B. De Lotto, E. de Oña Wilhelmi, C. Delgado Mendez, M. Doert, D. Dominis Prester, D. Dorner, M. Doro, S. Einecke, D. Eisenacher, D. Elsaesser, M. V. Fonseca, L. Font, K. Frantzen, C. Fruck, D. Galindo, R. J. García López, M. Garczarczyk, D. Garrido Terrats, M. Gaug, N. Godinović, A. González Muñoz, S. R. Gozzini, D.

- Hadasch, Y. Hanabata, M. Hayashida, J. Herrera, D. Hildebrand, J. Hose, D. Hrupec, W. Idec, V. Kadenius, H. Kellermann, K. Kodani, Y. Konno, J. Krause, H. Kubo, J. Kushida, A. La Barbera, D. Lelas, N. Lewandowska, E. Lindfors, S. Lombardi, M. López, R. López-Coto, A. López-Oramas, E. Lorenz, I. Lozano, M. Makariev, K. Mallot, G. Maneva, N. Mankuzhiyil, K. Mannheim, L. Maraschi, B. Marcote, M. Mariotti, M. Martínez, D. Mazin, U. Menzel, J. M. Miranda, R. Mirzoyan, A. Moralejo, P. Munar-Adrover, D. Nakajima, A. Niedzwiecki, K. Nilsson, K. Nishijima, K. Noda, N. Nowak, R. Orito, A. Overkemping, S. Paiano, M. Palatiello, D. Paneque, R. Paoletti, J. M. Paredes, X. Paredes-Fortuny, M. Persic, P. G. Prada Moroni, E. Prandini, S. Preziuso, I. Puljak, R. Reinthal, W. Rhode, M. Ribó, J. Rico, J. Rodriguez Garcia, S. Rügamer, A. Saggion, T. Saito, K. Saito, K. Satalecka, V. Scalzotto, V. Scapin, C. Schultz, T. Schweizer, S. N. Shore, A. Sillanpää, J. Sitarek, I. Snidaric, D. Sobczynska, F. Spanier, V. Stamatescu, A. Stamerra, T. Steinbring, J. Storz, M. Strzys, L. Takalo, H. Takami, F. Tavecchio, P. Temnikov, T. Terzić, D. Tescaro, M. Teshima, J. Thaele, O. Tibolla, D. F. Torres, T. Toyama, A. Treves, M. Uellenbeck, P. Vogler, R. M. Wagner, R. Zanin, D. Horns, J. Martín, and M. Meyer, “Measurement of the Crab Nebula spectrum over three decades in energy with the MAGIC telescopes”, *Journal of High Energy Astrophysics* **5** (2015).
- [119] P. S. Iroshnikov, “Turbulence of a conducting fluid in a strong magnetic field”, **7** (1964).
- [120] R. H. Kraichnan, “Inertial-range spectrum of hydromagnetic turbulence”, *Physics of Fluids* **8** (1965).
- [121] C. Thompson and O. Blaes, “Magnetohydrodynamics in the extreme relativistic limit”, **57** (1998).
- [122] G. Gogoberidze and Y. M. Voitenko, “Spectrum of imbalanced Alfvénic turbulence at ion-kinetic scales in the solar wind”, **365** (2020).
- [123] J. Cho and H. Kim, “Spectral evolution of helical electron magnetohydrodynamic turbulence”, *Journal of Geophysical Research (Space Physics)* **121** (2016).
- [124] R. Meyrand, J. Squire, A. A. Schekochihin, and W. Dorland, “On the violation of the zeroth law of turbulence in space plasmas”, *Journal of Plasma Physics* **87** (2021).
- [125] D. Grošelj, A. Mallet, N. F. Loureiro, and F. Jenko, “Fully Kinetic Simulation of 3D Kinetic Alfvén Turbulence”, *\prl* **120** (2018).

- [126] R. Schlickeiser, “Cosmic-Ray Transport and Acceleration. I. Derivation of the Kinetic Equation and Application to Cosmic Rays in Static Cold Media”, *\apj* **336** (1989).
- [127] B. D. G. Chandran, “Scattering of energetic particles by anisotropic magnetohydrodynamic turbulence with a goldreich-sridhar power spectrum”, **85** (2000).
- [128] V. Zhdankin, D. A. Uzdensky, G. R. Werner, and M. C. Begelman, “Numerical investigation of kinetic turbulence in relativistic pair plasmas - I. Turbulence statistics”, **474** (2018).
- [129] J. M. TenBarge, G. G. Howes, W. Dorland, and G. W. Hammett, “An oscillating Langevin antenna for driving plasma turbulence simulations”, *Computer Physics Communications* **185** (2014).
- [130] W. M. Elsasser, “The hydromagnetic equations”, *Physical Review* **79** (1950).
- [131] S. Galtier, S. V. Nazarenko, A. C. Newell, and A. Pouquet, “A weak turbulence theory for incompressible magnetohydrodynamics”, *Journal of Plasma Physics* **63** (2000).
- [132] J. M. TenBarge, B. Ripperda, A. Chernoglazov, A. Bhattacharjee, J. F. Mahlmann, E. R. Most, J. Juno, Y. Yuan, and A. A. Philippov, “Weak Alfvénic turbulence in relativistic plasmas I: asymptotic solutions”, arXiv e-prints (2021).
- [133] B. Ripperda, J. F. Mahlmann, A. Chernoglazov, J. M. TenBarge, E. R. Most, J. Juno, Y. Yuan, A. A. Philippov, and A. Bhattacharjee, “Weak Alfvénic turbulence in relativistic plasmas II: Current sheets and dissipation”, arXiv e-prints (2021).
- [134] A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, E. Quataert, and T. Tatsuno, “Astrophysical Gyrokinetics: Kinetic and Fluid Turbulent Cascades in Magnetized Weakly Collisional Plasmas”, *\apjs* **182** (2009).
- [135] S. Boldyrev and J. C. Perez, “Spectrum of weak magnetohydrodynamic turbulence”, **103** (2009).
- [136] S. Boldyrev, J. C. Perez, J. E. Borovsky, and J. J. Podesta, “Spectral scaling laws in magnetohydrodynamic turbulence simulations and in the solar wind”, **741** (2011).
- [137] Y. Wang, S. Boldyrev, and J. C. Perez, “Residual energy in magnetohydrodynamic turbulence”, **740** (2011).

- [138] M. Takamoto and A. Lazarian, “Compressible Relativistic Magnetohydrodynamic Turbulence in Magnetically Dominated Plasmas and Implications for a Strong-coupling Regime”, [\apjl 831 \(2016\)](#).
- [139] Y. Lithwick and P. Goldreich, “Compressible Magnetohydrodynamic Turbulence in Interstellar Plasmas”, [\apj 562 \(2001\)](#).
- [140] E. P. Alves, J. Zrake, and F. Fiuza, “Efficient Nonthermal Particle Acceleration by the Kink Instability in Relativistic Jets”, [\prl 121 \(2018\)](#).
- [141] J. Davelaar, A. A. Philippov, O. Bromberg, and C. B. Singh, “Particle acceleration in kink-unstable jets”, [896 \(2020\)](#).
- [142] M. A. Riquelme, E. Quataert, P. Sharma, and A. Spitkovsky, “Local two-dimensional particle-in-cell simulations of the collisionless magnetorotational instability”, [755 \(2012\)](#).
- [143] M. Hoshino, “Particle acceleration during magnetorotational instability in a collisionless accretion disk”, [773 \(2013\)](#).
- [144] M. W. Kunz, J. M. Stone, and E. Quataert, “Magnetorotational turbulence and dynamo in a collisionless plasma”, [117 \(2016\)](#).
- [145] K. Wong, V. Zhdankin, D. A. Uzdensky, G. R. Werner, and M. C. Begelman, “First-principles Demonstration of Diffusive-advective Particle Acceleration in Kinetic Simulations of Relativistic Plasma Turbulence”, [\apjl 893 \(2020\)](#).
- [146] C. Demidem, M. Lemoine, and F. Casse, “Particle acceleration in relativistic turbulence: A theoretical appraisal”, [102 \(2020\)](#).
- [147] B. Teaca, M. S. Weidl, F. Jenko, and R. Schlickeiser, “Acceleration of particles in imbalanced magnetohydrodynamic turbulence”, [\pre 90 \(2014\)](#).
- [148] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gauthier, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, and N. Wilkins-Diehr, “XSEDE: Accelerating scientific discovery”, [Computing in Science & Engineering 16 \(2014\)](#).

- [149] M. L. McConnell, A. A. Zdziarski, K. Bennett, H. Bloemen, W. Collmar, W. Hermsen, L. Kuiper, W. Paciesas, B. F. Phlips, J. Poutanen, J. M. Ryan, V. Schonfelder, H. Steinle, and A. W. Strong, “The Soft Gamma-Ray Spectral Variability of Cygnus X-1”, en, *The Astrophysical Journal* **572** (2002).
- [150] A. A. Zdziarski and M. Gierliński, “Radiative processes, spectral states and variability of black-hole binaries”, *Progress of Theoretical Physics Supplement* **155** (2004).
- [151] A. A. Zdziarski, J. E. Grove, J. Poutanen, A. R. Rao, and S. V. Vadawale, “OSSE and RXTE observations of GRS 1915+105: Evidence for nonthermal comptonization”, **554** (2001).
- [152] M. Gierliński, A. A. Zdziarski, J. Poutanen, P. S. Coppi, K. Ebisawa, and W. N. Johnson, “Radiation mechanisms and geometry of Cygnus X-1 in the soft state”, **309** (1999).
- [153] P. J. Armitage, C. S. Reynolds, and J. Chiang, “Simulations of accretion flows crossing the last stable orbit”, **548** (2001).
- [154] R. Shafee, J. C. McKinney, R. Narayan, A. Tchekhovskoy, C. F. Gammie, and J. E. McClintock, “Three-dimensional simulations of magnetized thin accretion disks around black holes: Stress in the plunging region”, **687** (2008).
- [155] S. C. Noble, J. H. Krolik, and J. F. Hawley, “Direct Calculation of the Radiative Efficiency of an Accretion Disk Around a Black Hole”, *\apj* **692** (2009).
- [156] M. J. Avara, J. C. McKinney, and C. S. Reynolds, “Efficiency of thin magnetically arrested discs around black holes”, *\mnras* **462** (2016).
- [157] Y. Zhu, S. W. Davis, R. Narayan, A. K. Kulkarni, R. F. Penna, and J. E. McClintock, “The eye of the storm: light from the inner plunging region of black hole accretion discs”, **424** (2012).
- [158] J. Dexter and O. Blaes, “A model of the steep power-law spectra and high-frequency quasi-periodic oscillations in luminous black hole X-ray binaries”, *\mnras* **438** (2014).
- [159] L. Spitzer, **Physics of fully ionized gases**, tex.adsnote: Provided by the SAO/NASA Astrophysics Data System tex.adsurl: <https://ui.adsabs.harvard.edu/abs/1956pfig.book.....S> (Interscience Publishers, Inc., 1956).

- [160] S. Nayakshin and F. Melia, “Self-consistent fokker-planck treatment of particle distributions in astrophysical plasmas”, **114** (1998).
- [161] G. R. Blumenthal and R. J. Gould, “Bremsstrahlung, Synchrotron Radiation, and Compton Scattering of High-Energy Electrons Traversing Dilute Gases”, *Reviews of Modern Physics* **42** (1970).
- [162] J. Dexter and E. Agol, “A Fast New Public Code for Computing Photon Orbits in a Kerr Spacetime”, *\apj* **696** (2009).
- [163] D. N. Page and K. S. Thorne, “Disk-Accretion onto a Black Hole. Time-Averaged Structure of Accretion Disk”, *\apj* **191** (1974).
- [164] C. T. Cunningham, “The effects of redshifts and focusing on the spectrum of an accretion disk around a Kerr black hole.”, **202** (1975).
- [165] L.-X. Li, E. R. Zimmerman, R. Narayan, and J. E. McClintock, “Multitemperature blackbody spectrum of a thin accretion disk around a kerr black hole: Model computations and comparison with observations”, **157** (2005).
- [166] R. J. H. Dunn, R. P. Fender, E. G. Körding, T. Belloni, and C. Cabanac, “A global spectral study of black hole X-ray binaries”, *\mnras* **403** (2010).
- [167] T. Muñoz-Darias, M. Coriat, D. S. Plant, G. Ponti, R. P. Fender, and R. J. H. Dunn, “Inclination and relativistic effects in the outburst evolution of black hole transients”, **432** (2013).
- [168] J. E. Grove, W. N. Johnson, R. A. Kroeger, K. McNaron-Brown, J. G. Skibo, and B. F. Philips, “Gamma-Ray Spectral States of Galactic Black Hole Candidates”, *\apj* **500** (1998).
- [169] J. Rodriguez, V. Grinberg, P. Laurent, M. Cadolle Bel, K. Pottschmidt, G. Pooley, A. Bodaghee, J. Wilms, and C. Gouiffès, “Spectral state dependence of the 0.4-2 MeV polarized emission in cygnus X-1 seen with INTEGRAL/IBIS, and links with the AMI radio data”, **807** (2015).
- [170] M. L. McConnell, J. M. Ryan, W. Collmar, V. Schönfelder, H. Steinle, A. W. Strong, H. Bloemen, W. Hermsen, L. Kuiper, K. Bennett, B. F. Philips, and J. C. Ling, “A High-Sensitivity Measurement of the MeV Gamma-Ray Spectrum of Cygnus X-1”, *\apj* **543** (2000).

- [171] B. Ripperda, M. Liska, K. Chatterjee, G. Musoke, A. A. Philippov, S. B. Markoff, A. Tchekhovskoy, and Z. Younsi, “Black Hole Flares: Ejection of Accreted Magnetic Flux through 3D Plasmoid-mediated Reconnection”, *\apjl* **924** (2022).
- [172] N. Scepi, J. Dexter, and M. C. Begelman, “Sgr A* X-ray flares from non-thermal particle acceleration in a magnetically arrested disc”, **511** (2022).
- [173] L. Comisso, E. Sobacchi, and L. Sironi, “Hard Synchrotron Spectra from Magnetically Dominated Plasma Turbulence”, *\apjl* **895** (2020).
- [174] S. W. Davis and I. Hubeny, “A Grid of Relativistic, Non-LTE Accretion Disk Models for Spectral Fitting of Black Hole Binaries”, *\apjs* **164** (2006).
- [175] G. Wardziński and A. A. Zdziarski, “Thermal synchrotron radiation and its Comptonization in compact X-ray sources”, **314** (2000).
- [176] G. Wardziński and A. A. Zdziarski, “Erratum: Effects of non-thermal tails in Maxwellian electron distributions on synchrotron and Compton processes”, *\mnras* **327** (2001).
- [177] J. M. Mehlhaff, G. R. Werner, D. A. Uzdensky, and M. C. Begelman, “Pair-regulated Klein-Nishina relativistic magnetic reconnection with applications to blazars and accreting black holes”, **508** (2021).
- [178] C. F. Gammie, J. C. McKinney, and G. Tóth, “HARM: A numerical scheme for general relativistic magnetohydrodynamics”, **589** (2003).
- [179] M. Mościbrodzka, C. F. Gammie, J. C. Dolence, H. Shiokawa, and P. K. Leung, “Radiative Models of SGR A* from GRMHD Simulations”, *The Astrophysical Journal* **706** (2009).
- [180] M. Mościbrodzka, H. Falcke, H. Shiokawa, and C. F. Gammie, “Observational appearance of inefficient accretion flows and jets in 3D GRMHD simulations: Application to Sagittarius A*”, *Astronomy and Astrophysics* **570** (2014).
- [181] G. R. Werner, D. A. Uzdensky, B. Cerutti, K. Nalewajko, and M. C. Begelman, “The extent of power-law energy spectra in collisionless relativistic magnetic reconnection in pair plasmas”, **816** (2016).

- [182] **Numerical Recipes: The Art of Scientific Computing**, 3rd (Cambridge University Press, 2007).
- [183] P. Goldreich and S. Tremaine, “Disk-satellite interactions.”, **241** (1980).
- [184] H. Bondi, “On spherically symmetrical accretion”, **112** (1952).
- [185] E. Lega, A. Crida, B. Bitsch, and A. Morbidelli, **440** (2014).
- [186] P. Benítez-Llambay, F. Masset, G. Koenigsberger, and J. Szulagyi, *Nature Letters* **520** (2015).
- [187] F. S. Masset, “Coorbital thermal torques on low-mass protoplanets”, **472** (2017).
- [188] A. Gruzinov, Y. Levin, and C. D. Matzner, “Negative dynamical friction on compact objects moving through dense gas”, **492** (2020).
- [189] H. Eklund and F. S. Masset, “Evolution of eccentricity and inclination of hot protoplanets embedded in radiative discs”, **469** (2017).
- [190] W. Kley and R. P. Nelson, “Planet-disk interaction and orbital evolution”, **50** (2012).
- [191] W. R. Ward, “Protoplanet migration by nebula tides”, **126** (1997).
- [192] S. .-. Paardekooper, C. Baruteau, and W. Kley, “A torque formula for non-isothermal Type I planetary migration - II. Effects of diffusion”, **410** (2011).
- [193] C. P. McNally, R. P. Nelson, S.-J. Paardekooper, P. Benítez-Llambay, and O. Gressel, “Low-mass planet migration in three-dimensional wind-driven inviscid discs: a negative corotation torque”, **493** (2020).
- [194] O. Chrenko, M. Brož, and M. Lambrechts, “Eccentricity excitation and merging of planetary embryos heated by pebble accretion”, **606** (2017).
- [195] C. W. Ormel and H. H. Klahr, “The effect of gas drag on the growth of protoplanets. Analytical expressions for the accretion of small bodies in laminar disks”, **520** (2010).
- [196] M. Lambrechts and A. Johansen, “Rapid growth of gas-giant cores by pebble accretion”, **544** (2012).
- [197] O. M. Guilera, N. Cuello, M. Montesinos, M. M. Miller Bertolami, M. P. Ronco, J. Cuadra, and F. S. Masset, “Thermal torque effects on the migration of growing low-mass planets”, **486** (2019).

- [198] T. Baumann and B. Bitsch, “Influence of migration models and thermal torque on planetary growth in the pebble accretion scenario”, **637** (2020).
- [199] D. Syer, C. J. Clarke, and M. J. Rees, “Star-disc interactions near a massive black hole”, **250** (1991).
- [200] I. Shlosman and M. C. Begelman, “Self-gravitating accretion disks in active galactic nuclei”, **329** (1987).
- [201] J. Goodman, “Self-gravity and quasi-stellar object discs”, **339** (2003).
- [202] Y. Levin, “Starbursts near supermassive black holes: young stars in the Galactic Centre, and gravitational waves in LISA band”, **374** (2007).
- [203] A. J. Dittmann and M. C. Miller, “Star formation in accretion discs and SMBH growth”, **493** (2020).
- [204] J. Bellovary, M. Mac Low, B. McKernan, and K. Ford, **819** (2016).
- [205] A. Secunda, J. Bellovary, M. Mac Low, K. Ford, N. Leigh, W. Lyra, and Z. Sandor, **878** (2019).
- [206] H. Tagawa, Z. Haiman, and B. Kocsis, “Formation and evolution of compact-object binaries in AGN disks”, **898** (2020).
- [207] N. C. Stone, B. D. Metzger, and Z. Haiman, “Assisted inspirals of stellar mass black holes embedded in AGN discs: solving the ‘final au problem’”, **464** (2017).
- [208] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, K. Agatsuma, and N. Aggarwal, “Binary black hole population properties inferred from the first and second observing runs of advanced LIGO and advanced virgo”, **882** (2019).
- [209] A. M. Derdzinski, D. D’Orazio, P. Duffell, Z. Haiman, and A. MacFadyen, “Probing gas disc physics with LISA: simulations of an intermediate mass ratio inspiral in an accretion disc”, **486** (2019).
- [210] D. A. Velasco Romero and F. S. Masset, “Numerical study of dynamical friction with thermal effects - I. Comparison to linear theory”, **483** (2019).
- [211] D. A. Velasco Romero and F. S. Masset, “Dynamical friction with radiative feedback - II. High-resolution study of the subsonic regime”, **495** (2020).

- [212] S. Fromenteau and F. S. Masset, “Impact of thermal effects on the evolution of eccentricity and inclination of low-mass planets”, [485 \(2019\)](#).
- [213] O. Chrenko and M. Lambrechts, “Oscillatory migration of accreting protoplanets driven by a 3D distortion of the gas flow”, [626 \(2019\)](#).
- [214] C. J. White, J. M. Stone, and C. F. Gammie, “An extension of the Athena++ code framework for GRMHD based on advanced riemann solvers and staggered-mesh constrained transport”, [225 \(2016\)](#).
- [215] K. G. Felker and J. M. Stone, “A fourth-order accurate finite volume method for ideal MHD via upwind constrained transport”, [Journal of Computational Physics 375 \(2018\)](#).
- [216] F. S. Masset and J. C. B. Papaloizou, “Runaway migration and the formation of hot jupiters”, [588 \(2003\)](#).
- [217] S. .-. Paardekooper, “Dynamical corotation torques on low-mass planets”, [444 \(2014\)](#).
- [218] M. A. Jiménez and F. S. Masset, “Improved torque formula for low- and intermediate-mass planetary migration”, [471 \(2017\)](#).
- [219] X. Li, P. Chang, Y. Levin, C. D. Matzner, and P. J. Armitage, “Simulation of a compact object with outflows moving through a gaseous background”, [494 \(2020\)](#).
- [220] C. Misner, K. Throne, and J. Wheeler, **Gravitation** (Princeton University Press).
- [221] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, “Rotating Black Holes: Locally Nonrotating Frames, Energy Extraction, and Scalar Synchrotron Radiation”, [The Astrophysical Journal 178 \(1972\)](#).
- [222] R. Penrose, “Gravitational Collapse: the Role of General Relativity”, [Nuovo Cimento Rivista Serie 1 \(1969\)](#).
- [223] R. D. Blandford and R. L. Znajek, “Electromagnetic extraction of energy from Kerr black holes.”, [Monthly Notices of the Royal Astronomical Society 179 \(1977\)](#).
- [224] B. Lowell, J. Jacquemin-Ide, A. Tchekhovskoy, and A. Duncan, **Rapid Black Hole Spin-down by Thick Magnetically Arrested Disks**, tech. rep., Publication Title: arXiv e-prints ADS Bibcode: 2023arXiv230201351L Type: article (Feb. 2023).

- [225] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczkó, D. Ball, M. Baloković, J. Barrett, D. Bintley, L. Blackburn, W. Boland, K. L. Bouman, G. C. Bower, M. Bremer, C. D. Brinkerink, R. Brissenden, S. Britzen, A. E. Broderick, D. Brogiere, T. Bronzwaer, D.-Y. Byun, J. E. Carlstrom, A. Chael, C.-k. Chan, S. Chatterjee, K. Chatterjee, M.-T. Chen, Y. Chen, I. Cho, P. Christian, J. E. Conway, J. M. Cordes, G. B. Crew, Y. Cui, J. Davelaar, M. De Laurentis, R. Deane, J. Dempsey, G. Desvignes, J. Dexter, S. S. Doeleman, R. P. Eatough, H. Falcke, V. L. Fish, E. Fomalont, R. Fraga-Encinas, W. T. Freeman, P. Friberg, C. M. Fromm, J. L. Gómez, P. Galison, C. F. Gammie, R. García, O. Gentaz, B. Georgiev, C. Goddi, R. Gold, M. Gu, M. Gurwell, K. Hada, M. H. Hecht, R. Hesper, L. C. Ho, P. Ho, M. Honma, C.-W. L. Huang, L. Huang, D. H. Hughes, S. Ikeda, M. Inoue, S. Issaoun, D. J. James, B. T. Jannuzi, M. Janssen, B. Jeter, W. Jiang, M. D. Johnson, S. Jorstad, T. Jung, M. Karami, R. Karuppusamy, T. Kawashima, G. K. Keating, M. Kettenis, J.-Y. Kim, J. Kim, J. Kim, M. Kino, J. Y. Koay, P. M. Koch, S. Koyama, M. Kramer, C. Kramer, T. P. Krichbaum, C.-Y. Kuo, T. R. Lauer, S.-S. Lee, Y.-R. Li, Z. Li, M. Lindqvist, K. Liu, E. Liuzzo, W.-P. Lo, A. P. Lobanov, L. Loinard, C. Lonsdale, R.-S. Lu, N. R. MacDonald, J. Mao, S. Markoff, D. P. Marrone, A. P. Marscher, I. Martí-Vidal, S. Matsushita, L. D. Matthews, L. Medeiros, K. M. Menten, Y. Mizuno, I. Mizuno, J. M. Moran, K. Moriyama, M. Moscibrodzka, C. Müller, H. Nagai, N. M. Nagar, M. Nakamura, R. Narayan, G. Narayanan, I. Natarajan, R. Neri, C. Ni, A. Noutsos, H. Okino, H. Olivares, G. N. Ortiz-León, T. Oyama, F. Özel, D. C. M. Palumbo, N. Patel, U.-L. Pen, D. W. Pesce, V. Piétu, R. Plambeck, A. PopStefanija, O. Porth, B. Prather, J. A. Preciado-López, D. Psaltis, H.-Y. Pu, V. Ramakrishnan, R. Rao, M. G. Rawlings, A. W. Raymond, L. Rezzolla, B. Ripperda, F. Roelofs, A. Rogers, E. Ros, M. Rose, A. Roshanineshat, H. Rottmann, A. L. Roy, C. Ruszczyk, B. R. Ryan, K. L. J. Rygl, S. Sánchez, D. Sánchez-Arguelles, M. Sasada, T. Savolainen, F. P. Schloerb, K.-F. Schuster, L. Shao, Z. Shen, D. Small, B. W. Sohn, J. SooHoo, F. Tazaki, P. Tiede, R. P. J. Tilanus, M. Titus, K. Toma, P. Torne, T. Trent, S. Trippe, S. Tsuda, I. van Bemmell, H. J. van Langevelde, D. R. van Rossum, J. Wagner, J. Wardle, J. Weintraub, N. Wex, R. Wharton, M. Wielgus, G. N. Wong, Q. Wu, K. Young, A. Young, Z. Younsi, F. Yuan, Y.-F. Yuan, J. A. Zensus, G. Zhao, S.-S. Zhao, Z. Zhu, J.-C. Algaba, A. Allardi, R. Amestica, J. Anczarski,

U. Bach, F. K. Baganoff, C. Beaudoin, B. A. Benson, R. Berthold, J. M. Blanchard, R. Blundell, S. Bustamente, R. Cappallo, E. Castillo-Domínguez, C.-C. Chang, S.-H. Chang, S.-C. Chang, C.-C. Chen, R. Chilson, T. C. Chuter, R. Córdova Rosado, I. M. Coulson, T. M. Crawford, J. Crowley, J. David, M. Derome, M. Dexter, S. Dornbusch, K. A. Dudevoir, S. A. Dzib, A. Eckart, C. Eckert, N. R. Erickson, W. B. Everett, A. Faber, J. R. Farah, V. Fath, T. W. Folkers, D. C. Forbes, R. Freund, A. I. Gómez-Ruiz, D. M. Gale, F. Gao, G. Geertsema, D. A. Graham, C. H. Greer, R. Grosslein, F. Gueth, D. Haggard, N. W. Halverson, C.-C. Han, K.-C. Han, J. Hao, Y. Hasegawa, J. W. Henning, A. Hernández-Gómez, R. Herrero-Illana, S. Heyminck, A. Hirota, J. Hoge, Y.-D. Huang, C. M. V. Impellizzeri, H. Jiang, A. Kamble, R. Keisler, K. Kimura, Y. Kono, D. Kubo, J. Kuroda, R. Lacasse, R. A. Laing, E. M. Leitch, C.-T. Li, L. C. .-. Lin, C.-T. Liu, K.-Y. Liu, L.-M. Lu, R. G. Marson, P. L. Martin-Cocher, K. D. Massingill, C. Matulonis, M. P. McColl, S. R. McWhirter, H. Messias, Z. Meyer-Zhao, D. Michalik, A. Montaña, W. Montgomerie, M. Mora-Klein, D. Muders, A. Nadolski, S. Navarro, J. Neilsen, C. H. Nguyen, H. Nishioka, T. Norton, M. A. Nowak, G. Nystrom, H. Ogawa, P. Oshiro, T. Oyama, H. Parsons, S. N. Paine, J. Peñalver, N. M. Phillips, M. Poirier, N. Pradel, R. A. Primiani, P. A. Raffin, A. S. Rahlin, G. Reiland, C. Risacher, I. Ruiz, A. F. Sáez-Madaín, R. Sassella, P. Schellart, P. Shaw, K. M. Silva, H. Shiokawa, D. R. Smith, W. Snow, K. Souccar, D. Sousa, T. K. Sridharan, R. Srinivasan, W. Stahm, A. A. Stark, K. Story, S. T. Timmer, L. Vertatschitsch, C. Walther, T.-S. Wei, N. Whitehorn, A. R. Whitney, D. P. Woody, J. G. A. Wouterloot, M. Wright, P. Yamaguchi, C.-Y. Yu, M. Zaballos, S. Zhang, and L. Ziurys, “First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole”, *The Astrophysical Journal* **875** (2019).

- [226] E.ourgoulhon, **3+1 Formalism and Bases of Numerical Relativity**, arXiv:gr-qc/0703035, Mar. 2007.

Appendix A

Local Simulations of Heating Torques on a Luminous Body in an Accretion Disk

Prologue

Objects such as stars and black holes can form in the thin accretion disk around a supermassive black hole. Because these objects are embedded in the gas of the accretion disk, they can accrete material and heat up the disk around themselves. Because these objects are not co-rotating with the disk, the torques present due to the heated gas are asymmetric and can cause the object to migrate within the AGN disk. In this paper, I examine the scaling of this torque with variables such as luminosity of the body and the thermal conductivity of the gas. I use hydrodynamic simulations in a shearing box, a local simulation that captures the shear flow but neglects vertical and radial structure of the accretion disk. This work was completed over the course of the 2019 summer school at the Center for Computational Astrophysics under the guidance of Phil Armitage and Yan-Fei Jiang. It was published in Ref. [7].

Abstract

A luminous body embedded in an accretion disk can generate asymmetric density perturbations that lead to a net torque and thus orbital migration of the body. Linear theory has shown that this heating torque gives rise to a migration term linear in the body's mass that can oppose or even reverse that arising from the sum of gravitational Lindblad and co-orbital torques. We use high-resolution local 3D shearing sheet simulations of a zero-mass test particle in an unstratified disk to assess the accuracy and domain of applicability of the linear theory. We find agreement between analytic and simulation results to better than 10% in the low luminosity, low thermal conductivity regime, but measure deviations in both the non-

linear (high luminosity) and the high thermal conductivity regimes. In the non-linear regime, linear theory overpredicts the acceleration due to the heating torque, potentially due to the neglect of non-linear terms in the heat flux. In the high thermal conductivity regime linear theory underpredicts the acceleration, which scales with a power-law index of -1 rather than $-3/2$, although here both non-linear and computational constraints play a role. We discuss the impact of the heating torque for the evolution of low-mass planets in protoplanetary disks, and for massive stars or accreting compact objects embedded in Active Galactic Nuclei disks. For the latter case, we show that the thermal torque is likely to be the dominant physical effect at disk radii where the optical depth drops below $\tau \lesssim 0.07\alpha^{-3/2}\epsilon c/v_K$.

A.1 Introduction

Planets, stars or compact objects orbiting within accretion disks perturb surrounding gas due to gravitational forces [183], accretion [184], the release of heat or radiation [185–187], and winds [188]. It is commonly the case that the resulting density perturbations leading and trailing the orbital motion are asymmetric, producing a gravitational back-reaction and a non-zero torque on the body. The torque leads to orbital migration—either an increase or a decrease in the semi-major axis—and evolution of any eccentricity or inclination (usually in the sense of damping, but in the opposite sense when combined with the release of heat [189]). In many circumstances of interest the time scale for migration is short compared to the disk lifetime, making it probable that observable properties of the system are substantially shaped by the effects of migration.

The longest-studied torque is that due to the purely gravitational perturbation of the disk-embedded object. It is made up of two independent components, one from waves excited at Lindblad resonances and one exerted in the co-orbital region [190], both of which scale as the square of the object’s mass. The net Lindblad torque (summing the opposite-signed contributions from interior and exterior resonances) has some dependence on disk properties, but is mostly due to intrinsic asymmetries in the interaction and almost always leads to inward migration [191]. The co-orbital torque, on the other hand, can lead to either inward or outward migration, and depends in a complex way on numerous properties of the disk [including radial gradients of vortensity and entropy, viscosity, thermal diffusivity, and disk winds; 192, 193].

Numerical simulations show that thermal effects, either in the disk gas in the vicinity of the planet or associated with the release of heat or radiation from a luminous body, result in additional torques [185, 186, 194]. Unlike the purely gravitational torques, thermal effects can (in principle) remain significant even for very low mass bodies. In particular, Ref. [187], using linear perturbation theory, identified a “heating torque” that arises when an orbiting body injects thermal energy into the surrounding disk. The thermal energy leads to the formation of low-density lobes near the planet, which are generically asymmetric, producing a torque. This heating torque is due purely to the injection of luminosity into the surrounding gas; the effect of the planet’s gravitational potential leads to another type of thermal torque called the “cold thermal torque”. The cold and the heating thermal torques can be separated and studied independently in the linear regime, that is, whereas the cold thermal torque is only present for a massive planet, the heating torque can be studied for a massless planet. Both the heating torque and the cold thermal torque can be on the same order of magnitude as other torques that cause migration (such as the Lindblad torque), and typically lead to outward migration.

The consequences of thermal torques on the migration rate of disk-embedded objects have been studied in the context of low-mass planet formation, where Lindblad torques alone would cause planets with masses of the order of the Earth’s mass to migrate toward the central star on a timescale shorter than the disk lifetime. The luminosity on these mass scales typically results from pebble accretion [195, 196]. The heating torque modifies the predicted map of where in the disk inward and outward migration occur [197], though the consequences for the final population of planets that form may be modest [198].

Heating torques could also impact the migration rate of luminous bodies such as stars and accreting stellar-mass black holes, which can be captured [199] or form [200–203] in the gas disks around supermassive black holes. Heating torques could interact with other gas torques (e.g. Lindblad torque, corotation torques) to form a migration trap—a radius in Active Galactic Nuclei (AGN) disk where the net torque is zero. Such migration traps would host an increased density of objects and provide a possible formation location for intermediate-mass black holes [204] or for stellar-mass black hole binaries [205, 206]. Stellar-mass binaries merging **within** an AGN disk could contribute to the observed LIGO population [207, 208], while stellar-mass black holes merging with the central supermassive black hole are future LISA sources, whose detailed properties may be modified by migration torques [209].

Heating torques have been studied analytically [187] and using global numerical simulations [185, 186, 194]. Here, we complement these prior studies using a local shearing box model for the disk. By simulating a luminous body in the limit where its mass goes to zero, using 32 zones per characteristic wavelength of the heating torque, we are able (a) to isolate the heating torque from the cold thermal torque and (b) to fully resolve the influence of the heating torque on the disk. Our work effectively extends the thorough numerical investigation of a luminous body travelling through a homogeneous medium [210, 211] to the case of a luminous body embedded within a shear flow. An important difference between these past studies and the present study is the massless nature of our planet, a piece of physics which we do not consider because the so-called “cold thermal torque” due to the gravitational potential of the planet should be separable from the heating torque in the linear regime [187]. The main questions we seek to answer are:

- (i) What are the numerical prerequisites needed to reproduce the [187] linear theory, and how accurate is that theory when the approximations involved are relaxed?
- (ii) When do non-linear effects set in, and how do they change the linear theory’s prediction for the thermal torque?
- (iii) Is the heating torque important for stars and accreting compact objects embedded within AGN disks?

The structure of the paper is as follows: we summarize the analytic theory for the heating torque resulting from a luminous body in a shear flow in §A.2.1 and describe our numerical methods in §A.2.2. Our numerical results are presented in §A.3. §A.4.1 discusses the limits to the analytic theory, and §A.4.2 discusses applications of the model to luminous objects in AGN and protoplanetary disks. We conclude in §A.5.

A.2 Methods

A.2.1 Analytic Results

Analysis of the local hydrodynamic equations with thermal conductivity shows that the effects of a massive body’s gravitational potential and its luminosity on a surrounding disk can be separated and

studied independently in the linear regime [187, Eq. 34]. We take advantage of this separation to focus solely on the “heating torque”, the torque due to the density perturbation that is sourced by thermal energy diffusing outward from a luminous body, though in the presence of orbital eccentricities and inclinations the applicability of the linear regime is limited [189, 194, 212]. Fig. A.1 illustrates how the asymmetry in this perturbation, resulting from the displacement of the orbiting body from co-rotation, leads to a net torque. To aid in the interpretation of our numerical results, we summarize the key assumptions and results from Ref. [187].

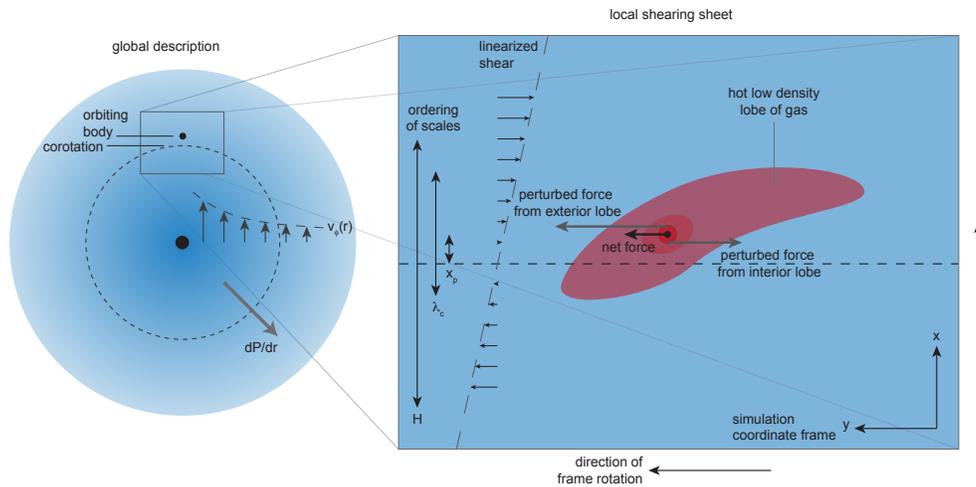


Figure A.1: Illustration of the physics leading to a heating torque (the gravitational potential of the body is neglected). Heat diffusing away from a luminous, disk-embedded body, is sheared by the sub-Keplerian disk flow, forming hot low density lobes. These lobes are asymmetric interior to / exterior to the body, because the body is displaced from exact co-rotation due to the presence of a pressure gradient in the disk gas. The gravitational back-reaction from the heated lobes exerts a positive torque on the orbiting body, as illustrated by the vectors showing the *perturbed* force in the azimuthal direction, i.e. the difference between the force on a luminous massless body and the force on a non-luminous massless body. Other components of the force (e.g. in the radial direction) are not drawn because they are opposite and equal and thus do not contribute to the total torque.

Ref. [187] linearizes the hydrodynamic equations (see eqs. (A.9) to (A.11)), assuming a steady-state, in a local (“shearing box”) frame that co-rotates with the orbiting body. In this local frame, x corresponds to the radial direction, y to the azimuthal, and z to the vertical direction (perpendicular to the disk midplane) as illustrated in Fig. A.1. If there is a radial pressure gradient in the disk, there is an offset x_p between the

orbiting body and disk gas that has the same orbital velocity. This distance from co-rotation is given by,

$$x_p = -\frac{\partial_x p_0}{2q\Omega_0^2\rho_0}. \quad (\text{A.1})$$

Here p_0 (ρ_0) is the equilibrium background pressure (density), q is the shearing parameter (equal to 3/2 for Keplerian disks), and Ω_0 is the angular velocity of the local frame. For typical pressure profiles that decrease as a function of radius x_p is positive, implying that the body will sit further away from the central body than the gas rotating at the same angular velocity, experiencing a headwind.

Three characteristic scales enter the problem: the distance from co-rotation x_p , the characteristic size of the density perturbation caused by the body's luminosity λ_c , and the pressure scale height of the disk H . In the linear calculation it is assumed that the following hierarchy holds,

$$x_p \ll \lambda_c, \quad (\text{A.2})$$

$$\lambda_c \ll H. \quad (\text{A.3})$$

We refer to the first requirement for scale separation (Eq. A.2) as Assumption II and the second (Eq. A.3) as Assumption III (the first assumption is that of linearity). Assumption III allows the vertical density gradient of the local box to be neglected (justifying our use of unstratified simulations, although stratification can cause oscillatory torques when the opacity is not constant; [213]), while the small parameter associated with Assumption II is used extensively to expand the expected gravitational force from the under-density caused by the body's luminosity. The relative importance of these two hierarchies and the validity of the predicted net force on the body (Eq. A.5) is explored in § A.4.1.1.

The characteristic size of the disturbance and the net azimuthal force experienced by the body as a result of the heating torque mechanism are predicted to be [187, Eq. 83, and Eq. 109],

$$\lambda_c = 2\pi k_c^{-1} = 2\pi \sqrt{\frac{\chi}{q\Omega_0\gamma}}, \quad (\text{A.4})$$

$$F_y = \frac{0.322x_p\gamma^{3/2}(\gamma-1)GMLq^{1/2}\Omega_0^{1/2}}{\chi^{3/2}c_s^2}. \quad (\text{A.5})$$

Here γ is the adiabatic index ($\gamma = 5/3$ in all of the following simulations), $c_s^2 = \gamma p_0/\rho_0$ is the equilibrium sound speed, L is the luminosity emitted by the body, χ is the disk's thermal conductivity, and M is the mass

of the body. Crucially, Eq. A.5 is **linear** in the mass of the body. This feature of the heating torque allows us to calculate the force per unit mass (i.e. the body's acceleration) without needing to explicitly include the body's mass at all in the simulations.

The heating torque is of interest because it can be the same order of magnitude as the other torques in the system (such as the Lindblad torque). Defining

$$L_c = \frac{4\pi GM\chi\rho_0}{\gamma}, \quad (\text{A.6})$$

the heating torque can be written as [187, Eq. 144]

$$\Gamma^{heat} = 1.61 \frac{\gamma - 1}{\gamma} \frac{x_p}{\lambda_c/2\pi} \frac{L}{L_c} \Gamma_0, \quad (\text{A.7})$$

where

$$\Gamma_0 = \sqrt{2\pi}\rho_0 H r_0^4 \Omega_0^2 \left(\frac{M}{M_*}\right)^2 \left(\frac{r_0}{H}\right)^3, \quad (\text{A.8})$$

is of the order of the Lindblad torque. Here r_0 is the semi-major axis of the body, M_* is the mass of the central object, and H is the pressure scale height of the disk. Note that this definition of Γ_0 differs from the more widespread definition (in for instance [192]) by a factor of r_0/H .

In summary, the formula given by Eq. A.5 for the net force on a body (due to the asymmetric gravitational forces caused by the body's luminosity distributed by differential rotation) is predicted to hold under three conditions:

- Assumption I: perturbations of density and pressure should be much less than equilibrium values (linearity, $\rho' \ll \rho_0$).
- Assumption II: the offset from corotation x_p should be much less than the size of the disturbance λ_c ($x_p \ll \lambda_c$).
- Assumption III: the disturbance should be much smaller than the pressure scale height of the disk ($\lambda_c \ll H$).

In this work, we test the validity of the linear theory when one or more of these assumptions is violated.

A.2.2 Simulations

The linear theory is developed in the local “shearing sheet” approximation [187], which translates directly into a well-studied numerical set-up. We solve the inviscid hydrodynamic equations in a local approximation of a Cartesian box rotating around a massive body (a star or black hole, for instance) with orbital frequency Ω_0 , and add a source term to the energy density equation to model the luminosity. With ρ the mass density, e the energy density, P the pressure, and \mathbf{V} the velocity, the equations read,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{V}) = 0, \quad (\text{A.9})$$

$$\frac{\partial \rho \mathbf{V}}{\partial t} + \nabla \cdot (\rho \mathbf{V} \mathbf{V} + P \mathbf{I}) = -\rho \nabla \Phi_t - 2\rho \Omega \hat{z} \times \mathbf{V}, \quad (\text{A.10})$$

$$\frac{\partial e}{\partial t} + \nabla \cdot [\mathbf{V}(e + P) + F_H] = L \delta(x - x_p), \quad (\text{A.11})$$

where $F_H = -\chi \rho \nabla (e/\rho)$ is the heat flux and L is the total luminosity emitted by the body and $\delta(x - x_p)$ is the Dirac delta function. The gas has an adiabatic equation of state. $\Phi_t = -q \Omega_0^2 (x - x_p)^2$ is the tidal potential due to the central object. The vertical density gradient is neglected, both for consistency with Ref. [187] and for the same physical reasons discussed there, and the radial density gradient is modelled through a non-zero offset from co-rotation (it is neglected in the shearing-box, as justified by assuming that the background pressure does not change significantly over the short radial scales under consideration). The shearing parameter q is equal to 3/2 for the Keplerian flows studied in this work.

The Athena++ code is used to solve the above equations in the luminous body’s rest frame on a uniform Cartesian mesh [214, 215, Stone et al., 2020, submitted] with the Harten-Lax-van Leer-Contact Riemann solver. The simulation’s origin sits at the radial location where the gas orbits at the same frequency as the luminous body, such that the position of the body is fixed over the course of the simulation (10 orbits). We discuss the consequences of neglecting the radial motion of the body in response to the generated torque in §A.4.2.2. Both the origin and the body sit at the mid-plane of the disk ($z = 0$). In the fiducial run L1K1, the simulation domain spans $[4.13, 12.4, 4.13]H$ in the x (radial), y (azimuthal), and z (vertical) directions respectively, where H is the pressure scale height of the disk (defined through $H = c_s/\Omega_0$). The fiducial run L1K1 has a resolution of $[256, 192, 256]$ cells (i.e. $[62.0, 20.7, 62.0]$ cells/ H) and a value $x_p = 0.097 H$. Convergence with resolution and domain size is studied in §A.3.5.

The body's luminosity is modelled by directly injecting internal energy density into the gas via the energy density equation (A.11). In the analytic theory the injection term is $L \delta(x - x_p)$ (as in Eq. A.11). Since injection at a single point is not possible numerically, we implement this term in the simulations by adding an energy density $\ell_v \times \Delta t$ at each time to each cell whose center lies within an injection radius r_{rad} . Here, ℓ_v is the (constant) luminosity per volume and Δt is the time step as determined by the Courant condition. The total luminosity L injected at each time step can be calculated as $L = \ell_v \times n \times v$, where n is the number of cells included in the injection region and v is each cell's volume; thus the total luminosity L depends on both the luminosity per volume ℓ_v and the injection radius r_{rad} . Unless otherwise specified, the injection radius is set such that the luminosity is evenly distributed into the eight cells neighboring the body. The effect of a finite injection radius (rather than a strict Dirac delta function) is explored in §A.3.5. We note again that according to the linear theory the torque due to the gravitational potential of the body can be separated from the torque due to the body's luminosity [187]. In this work we model only the heating torque, and do not include the gravitational potential of the body.

As is standard for the shearing-box set-up, all simulations use periodic boundary conditions in the azimuthal and vertical directions and shearing-periodic boundary conditions in the radial direction, the effects of which are discussed in §A.3.5. Generic units H (disk scale height) for length, Ω^{-1} ($2\pi/\Omega$ is one orbit) for time, and P_0 (background pressure) for energy density are used. The sound speed thus has units of $H\Omega$, acceleration has units of $H\Omega^2$, thermal diffusivity has units of $H^2\Omega$, and total injected luminosity has units of $P_0H^3\Omega$. These values can be scaled to various astrophysical systems, which are discussed in §A.4.

A.2.3 Diagnostics

Linear theory provides a prediction for the total gravitational force F_y experienced by the orbiting body as a result of the perturbed gas density (purely from the body's luminosity, not its gravity). We calculate this azimuthal force per unit body mass in the simulation on spherical shells by calculating the distance between the body and each cell, assuming that all of the cell's mass is located at its center, and using the y-component of the inverse square law with $GM = 1$ in code-units. The resulting acceleration can be plotted as a function of radius or summed over radius to directly compare to Eq. A.5. To avoid introducing artificial asymmetry

to the force (i.e. a systematically larger force on the $x < 0$ side), the summation stops at the shortest radius that fits inside the simulation domain in all directions. The limiting radius is thus $L_x - x_p$, where L_x is the half-width of the box in the x (radial) direction and x_p is the body's distance from corotation. This restriction is not expected to change the results significantly since the excluded portions are not a large fraction of the box, the gas there is not as perturbed, and the force from the gas there is attenuated by the inverse square of the radius).

The fractional change \tilde{f} in a quantity f is useful to establish the linearity of density perturbations,

$$\tilde{f} = \frac{f(t) - f(t=0)}{f(t=0)} = \frac{f'(t)}{f(t=0)}, \quad (\text{A.12})$$

where $f'(t) = f(t) - f(t=0)$ is the perturbation from equilibrium. We describe a simulation as being in the linear regime if the deviation from equilibrium values is no more than 5%.

To facilitate direct comparison with previous work (specifically, Ref. [187]'s Fig. 1), we calculate the perturbation in surface density σ' as the $k_z = 0$ mode of the Fourier transform $\hat{\rho}(x, y, k_z)$ of the density perturbation ρ' in the z -direction, i.e.

$$\hat{\rho}(x, y, k_z) = \int_{-\infty}^{\infty} \rho'(x, y, z) e^{-ik_z z} dz, \quad (\text{A.13})$$

$$\sigma(x, y) = \hat{\rho}(x/k_c, y/k_c, 0). \quad (\text{A.14})$$

Separating into the effect due to zero offset from corotation ($\sigma^{(0)}$) and the first order effect due to nonzero offset ($\sigma^{(1)}$), we find [187, Eq. 114]

$$\sigma'(x/k_c, y/k_c) = \sigma'^{(0)}(x/k_c, y/k_c) + x_p \sigma'^{(1)}(x/k_c, y/k_c). \quad (\text{A.15})$$

Note that σ' is the *perturbation* of the surface density; the unperturbed surface density σ is constant. The expected offset of $\sigma'^{(1)}$ is $x_p k_c = 0.59$ relative to Ref. [187]'s Fig. 1b.

A.3 Results

A.3.1 The Linear Regime

Fig. A.2 shows the mid-plane pressure and density perturbations derived from simulations in the linear and non-linear regimes of luminosity injection. Values of $\ell_v = 1.42 P_0 \Omega$ (Fig. A.2 left column;

physical values are discussed in A.4.2.2) lead to perturbations that are less than 5% of the equilibrium values, which we take to be in the linear regime. The measured pressure perturbations are two orders of magnitude smaller than the density perturbations. This is consistent with Ref. [187]’s estimation (Eq. 36) that $P' \ll H^2 \Omega_0^2 \rho' \approx c_s^2 \rho'$. With the value $c_s^2 = \gamma P / \rho = 1.00 H^2 \Omega^2$, P' should be much less than ρ' . The non-linear regime is illustrated in the right column of Fig. A.2, which injects two orders of magnitude more energy per timestep ($\ell_v = 142 P_0 \Omega$; simulation L100K1). The qualitative appearance of the pressure and density perturbations remain similar for this much higher rate of energy injection. Both simulations, as expected, quickly reach an equilibrium within approximately two orbits.

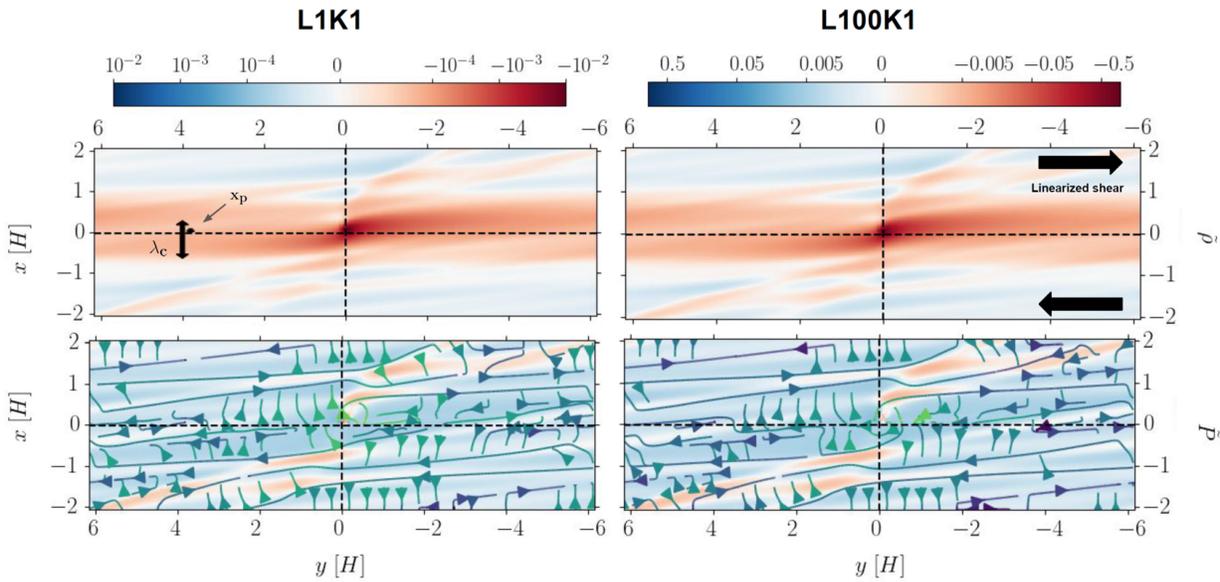


Figure A.2: Slices in the $z = 0$ plane of simulations with $\chi = 0.017 H^2 \Omega$ at $t = 5.0$ orbits. Top row: density perturbation as a percentage of initial (equilibrium) condition. Bottom row: perturbation in pressure as a percentage of initial (equilibrium) condition. Left column: $\ell_v = 1.42 P_0 \Omega$ (fiducial simulation L1K1, linear regime). Right column: $\ell_v = 142 P_0 \Omega$ (high luminosity simulation L100K1, non-linear regime). Note that (though similar in shape) the perturbations in L100K1 are larger in magnitude than the perturbations in L1K1. Velocity streamlines with the background shear profile subtracted are plotted in the lower panels, normalized to their maximum values for ease of comparison between L1K1 and L100K1. The distance from co-rotation x_p , the characteristic wavelength λ_c , and the direction of shear are marked with black arrows.

A.3.2 Net Azimuthal Acceleration as a Function of Radius

We take the $\ell_v = 1.42 P_0 \Omega$ run as our fiducial simulation L1K1 so as to be firmly in the linear regime. Using the technique described in §A.2.3, we plot as a function of radius the net gravitational force on the body per unit body mass as a result of the gas perturbed by the object’s luminosity. Fig. A.3 shows the result of summing up gas in front of and behind the body ($y > 0$ and $y < 0$, respectively), as well as the total force (green; right-hand scale). The bottom panel shows how the net azimuthal acceleration differs ahead/behind the body from the initial acceleration (which is non-zero on either side but which sums to zero) as well as how close the net azimuthal acceleration summed over all radii is to the linear prediction (horizontal lines). The noisiness of the data in Fig. A.3 is due to the fact that no interpolation was used in the calculation of the acceleration. Measurements from this simulation agree with the linear prediction to better than 10%.

A.3.3 Perturbation of the Surface Density

In addition to the value of the net force acting on the body, Ref. [187] calculates a map of the surface density perturbation predicted by the analytic theory. Fig. A.4 reproduces this map with simulation data (compare to Ref. [187]’s Fig. 1). The upper panel Fourier transforms the density perturbation of a simulation with no offset ($x_p = 0$), as outlined in §A.2.3. The lower panel Fourier transforms the density perturbation with the equilibrium value of x_p , then subtracts the zero offset case and divides by x_p to extract $\sigma'^{(1)}$. The perturbations are smaller in amplitude than expected, largely because the peak expected amplitude is very close to the luminous body and is not as resolved. However, the general shape of the perturbation agrees well with the linear prediction.

A.3.4 Scaling Relations

Linear theory predicts a linear dependence of the net gravitational force on the total luminosity L emitted by the body and a power-law dependence $F_y \propto \chi^{-3/2}$ on the thermal conductivity (Eq. A.5). To test these predictions we ran two suites of simulations: one that fixes the thermal conductivity and varies the total emitted luminosity, and one that fixes the total emitted luminosity and varies the thermal conductivity.

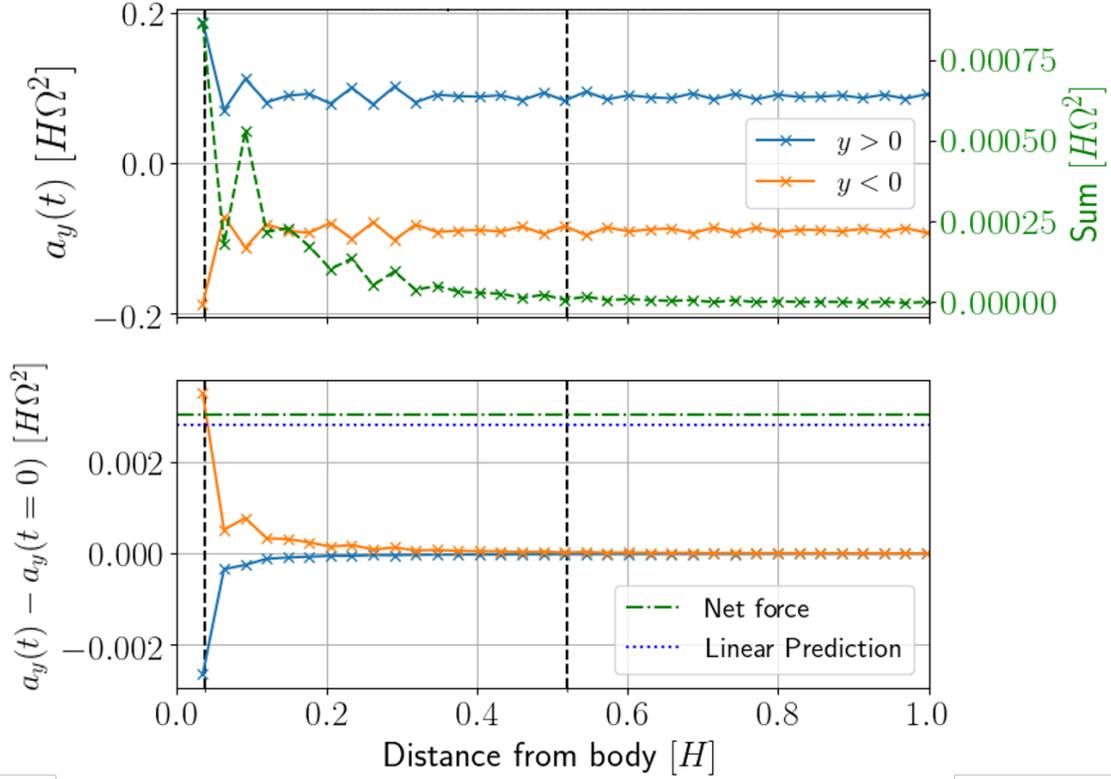


Figure A.3: Snapshot of gravitational acceleration on the body in the fiducial simulation L1K1 as a function of distance away from the body at $t = 5.0$ orbits. The top panel plots both the one-sided forces due to gas in front of ($y > 0$) and behind ($y < 0$) the body, as well as the sum of the forces (dashed green line; right scale). Vertical dashed lines show the size of the luminosity injection radius $r_{\text{rad}} = 0.04 H$ and half the characteristic wavelength $\lambda_c = 0.52 H$. The bottom panel shows the difference between the initial condition (which has net force equal to zero but one-sided forces on the order of the top panel's vertical values) and their values at five orbits. For reference the sum over all radii (with value $3.03 \times 10^{-3} H\Omega^2$) is plotted as a dash-dot horizontal line, and for comparison the linear theory's predicted value of $2.82 \times 10^{-3} H\Omega^2$ is plotted as a dotted horizontal line.

For the first suite we fixed $\chi = 0.017 H^2\Omega$ and varied ℓ_v over three orders of magnitude: from $\ell_v = 0.142 P_0\Omega$ to $142 P_0\Omega$. Simulations are considered to be in the linear regime if the perturbation never exceeds 5% of the equilibrium value. Fig. A.5 reveals a tight agreement with the linear prediction even an order of magnitude into the non-linear regime (indicated by green squares). As the injected luminosity increases even more, the linear theory begins to over-predict the measured force because of non-linear effects; this over-prediction is discussed in §A.4.1.2. In the linear regime at least, we are able to reproduce both the scaling and the normalization of the net force to within 10%.

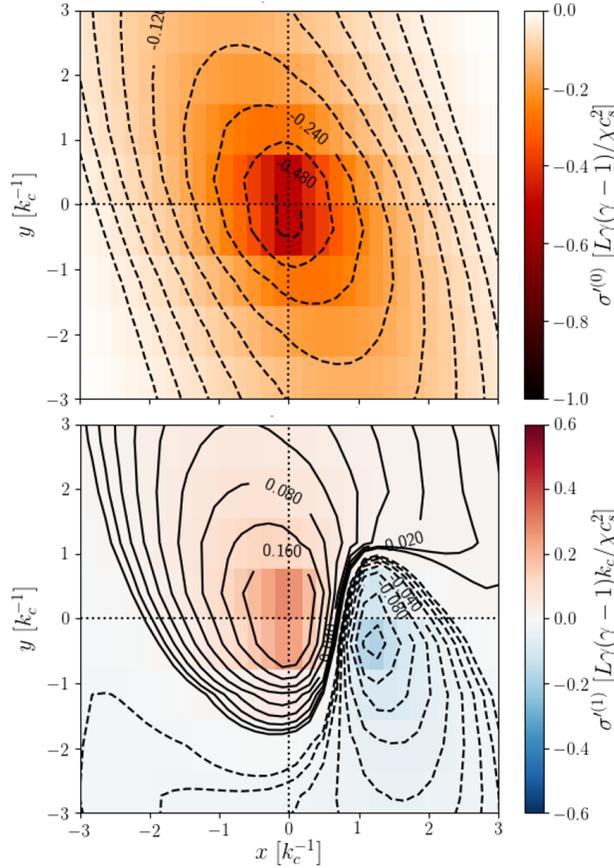


Figure A.4: Perturbation of surface density in units of $\gamma(\gamma - 1)L/\chi c_s^2$ due to the luminous body's heat at $t = 3.5$ orbits. Contour levels on the left are a geometric series with a ratio of $\sqrt{2}$ from -0.03 to -0.48 . On the right, contour levels have a ratio of 2 between them and run from ± 0.01 to ± 0.16 . Solid contour are positive values; dashed are negative. Thermal conductivity is $\chi = 0.017 H^2\Omega$, $\ell_v = 1.42P_0\Omega$. Similar to Ref. [187, Fig. 1].

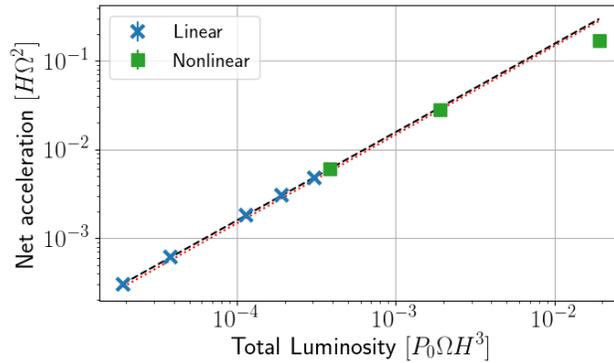


Figure A.5: Net azimuthal acceleration due to the gas's gravity as a function of total injected luminosity. Thermal conductivity is fixed at $\chi = 0.017 H^2\Omega$. The linear theory's prediction (slope: 1.0; red dotted line) matches the data well even into the non-linear regime, although the fit (slope: $1.0 \pm 6.5 \times 10^{-7}$; black dashed line) was determined using only the linear data points. Simulations are summarized in Table A.1.

	$L [P_0\Omega H^3]$	$\ell_v [P_0\Omega]$	$F_y [H\Omega^2]$ (Linear Prediction)	$F_y [H\Omega^2]$ (Measured)
L1K1	1.91×10^{-5}	0.142	2.82×10^{-4}	$3.05\times 10^{-4}\pm 9.76\times 10^{-7}$
	3.83×10^{-5}	0.284	5.64×10^{-4}	$6.10\times 10^{-4}\pm 1.95\times 10^{-6}$
	1.15×10^{-4}	0.853	1.69×10^{-3}	$1.83\times 10^{-3}\pm 5.75\times 10^{-6}$
	1.91×10^{-4}	1.42	2.82×10^{-3}	$3.03\times 10^{-3}\pm 9.46\times 10^{-6}$
*	3.06×10^{-4}	2.28	4.51×10^{-3}	$4.83\times 10^{-3}\pm 1.50\times 10^{-5}$
	3.83×10^{-4}	2.84	5.64×10^{-3}	$6.02\times 10^{-3}\pm 1.85\times 10^{-5}$
*	1.91×10^{-3}	14.2	2.82×10^{-2}	$2.83\times 10^{-2}\pm 8.02\times 10^{-5}$
* L100K1	1.91×10^{-2}	142	2.82×10^{-1}	$1.69\times 10^{-1}\pm 4.59\times 10^{-4}$

Table A.1: Summary of simulations used to calculate the scaling of net azimuthal acceleration with total injected luminosity L (Fig. A.5). All simulations have $\chi = 0.017 H^2\Omega$, corresponding to $\lambda_c = 0.52 H$, while all other parameters (e.g. resolution, offset of the body from corotation, described in the text) are that of the fiducial simulation L1K1. This set of simulations keeps λ_c/H and x_p/λ_c constant at 0.52 and 0.187, respectively. * indicates a simulation that has density fluctuations greater than 5% of the equilibrium value and has thus entered the non-linear regime. Simulation L1K1 is often referred to as the fiducial simulation, and L100K1 as the high luminosity simulation. Measured values are presented as the average between 1 and 10 orbits plus/minus one standard deviation of the value in time.

Assessing the validity of the analytic prediction for the scaling of the net force with thermal conductivity is substantially harder, because changing the conductivity also changes the characteristic wavelength λ_c . It is difficult to find a numerically tractable set of parameters that both (a) remains in the linear regime and (b) maintains the hierarchy of scales required by Ref. [187], over a substantial range in χ .

Fig. A.6 shows the measured dependence of the net force as a function of the thermal conductivity, at fixed luminosity. For sufficiently low values of the thermal conductivity, heat cannot diffuse away fast enough, causing the system to enter the non-linear regime (indicated by a green square). For high values of the thermal conductivity, the required scale separation $x_p \ll \lambda_c \ll H$ is lost (shown as orange dots). Only the blue crosses, at intermediate χ , remain linear and respect the scale hierarchy.

Fitting the data only at intermediate χ , we find that the dependence of net gravitational acceleration on conductivity is close to χ^{-1} , rather than the expected $\chi^{-3/2}$. We caution, however, that this fit is made over only a very limited range of χ . If, instead, we consider simulation data at higher values of χ , we find a dependence that appears to be closer to the analytically predicted power-law. The ideal regime for matching the linear theory appears to be around $\chi = 0.017 H^2 \Omega$, which in the simulations presented has ratios $x_p/\lambda_c = 0.187$ and $\lambda_c/H = 0.519$. Simulations using $\chi = 0.0061 H^2 \Omega$ to obtain ratios $x_p/\lambda_c = 0.311 = \lambda_c/H$ measured an acceleration lower than the linear prediction by a factor of two. This suggests that the requirement that $x_p \ll \lambda_c$ is more important for matching linear theory than $\lambda_c \ll H$. This hierarchy of the assumption is reasonable since the former is used in expanding the net force, whereas the latter is used to drop vertical density stratification [187]; see §A.4.1.1.

A.3.5 Numerical Considerations

In order to assess the robustness of the numerical results, we explored the dependence of the simulation results on domain size, resolution, boundary conditions, and injection radius. Of these factors, we find that the most important numerical effects are related to the size of the injection region. The analytic assumption that all the body's luminosity is deposited at a single point is both an approximation to the physical situation, and an idealization that cannot be achieved in grid-based numerical simulations. We find that for the fiducial parameters and a resolution that allows for an injection radius of $r_{\text{rad}} = 0.04 H$, the measured azimuthal force

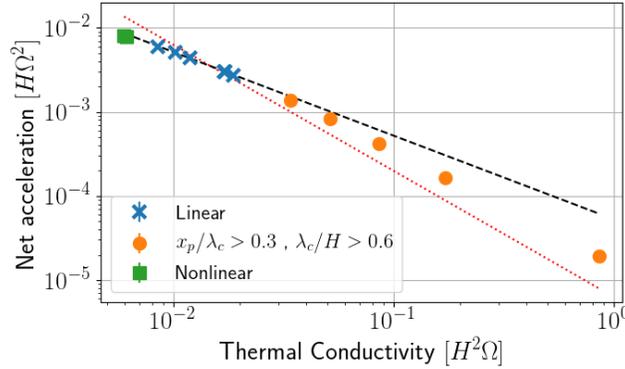


Figure A.6: Net azimuthal acceleration due to the gas’s gravity as a function of thermal conductivity. Total luminosity is fixed with $L = 1.96 \times 10^{-4} P_0 \Omega H^3$ ($\ell_v = 1.42 P_0 \Omega$). The linear theory (red dotted line) predicts a power-law index of -1.5 , whereas the fit (black dashed line) determines a power-law index of $-1.0 \pm 1.6 \times 10^{-4}$ and total luminosity $L = 6.7 \times 10^{-3} P_0 \Omega H^3$. The fit was determined solely from the simulations satisfying the hierarchies $x_p/\lambda_c < 0.3$ and $\lambda_c/H < 0.6$ (blue crosses). Simulations are summarized in Table A.2.

	$\chi [H^2\Omega]$	$\lambda_c/2 [H]$	$F_y [H\Omega^2]$ (Linear Prediction)	$F_y [H\Omega^2]$ (Measured)
*	5.97×10^{-3}	1.54×10^{-1}	1.36×10^{-2}	$7.96 \times 10^{-3} \pm 4.24 \times 10^{-5}$
*	6.14×10^{-3}	1.56×10^{-1}	1.31×10^{-2}	$7.79 \times 10^{-3} \pm 3.98 \times 10^{-5}$
	8.53×10^{-3}	1.84×10^{-1}	7.97×10^{-3}	$5.95 \times 10^{-3} \pm 2.61 \times 10^{-5}$
	1.02×10^{-2}	2.01×10^{-1}	6.07×10^{-3}	$5.06 \times 10^{-3} \pm 2.04 \times 10^{-5}$
	1.19×10^{-2}	2.17×10^{-1}	4.81×10^{-3}	$4.37 \times 10^{-3} \pm 1.64 \times 10^{-5}$
L1K1	1.71×10^{-2}	2.60×10^{-1}	2.82×10^{-3}	$3.03 \times 10^{-3} \pm 9.46 \times 10^{-6}$
	1.72×10^{-2}	2.61×10^{-1}	2.78×10^{-3}	$3.00 \times 10^{-3} \pm 9.46 \times 10^{-6}$
	1.88×10^{-2}	2.72×10^{-1}	2.44×10^{-3}	$2.74 \times 10^{-3} \pm 8.32 \times 10^{-6}$
	3.41×10^{-2}	3.67×10^{-1}	9.97×10^{-4}	$1.37 \times 10^{-3} \pm 3.54 \times 10^{-6}$
	5.12×10^{-2}	4.50×10^{-1}	5.43×10^{-4}	$8.24 \times 10^{-4} \pm 2.17 \times 10^{-6}$
	8.53×10^{-2}	5.80×10^{-1}	2.52×10^{-4}	$4.21 \times 10^{-4} \pm 1.35 \times 10^{-6}$
L1K10	1.71×10^{-1}	8.21×10^{-1}	8.91×10^{-5}	$1.65 \times 10^{-4} \pm 8.04 \times 10^{-7}$
L1K50	8.53×10^{-1}	1.84	7.97×10^{-6}	$1.93 \times 10^{-5} \pm 2.92 \times 10^{-7}$

Table A.2: Summary of simulations used to calculate the scaling of net azimuthal acceleration with conductivity χ (Fig. A.6). All simulations have $L = 1.9 \times 10^{-4} P_0 \Omega H^3$ ($\ell_v = 1.42 P_0 \Omega$), while all other parameters (e.g. resolution, offset of the body from corotation) are that of the fiducial simulation L1K1. The value of λ/H is easily read off; the value of x_p/λ_c is obtained by noting that $x_p = 0.097 H$. * indicates a simulation that has density fluctuations greater than 5% of the equilibrium value and has thus entered the non-linear regime. Simulation L1K1 is often referred to as the fiducial simulation, and L1K10 as the high conductivity simulation. Measured values are presented as the average between 1 and 10 orbits plus/minus one standard deviation of the value in time.

is 7.6% larger than the linear theory's prediction. Doubling the injection radius to $r_{\text{rad}} = 0.07 H$, at half the resolution, leads to an error with respect to the linear prediction of -25% , i.e. a decrease in resolution results in a measured azimuthal force smaller than the linear theory's prediction. An even higher resolution with a correspondingly small injection radius could result in even better agreement with the linear theory; however, at this point the question of more detailed physics close to the body would likely be more pressing.

To isolate the effect of changing spatial resolution from the effect of differing injection radii, we test for convergence with spatial resolution by keeping the same injection radius and changing the resolution. Due to the discretization of the region around the body, increasing the resolution will result in an injection region that closer approximates a sphere rather than a rectangular prism (as is the case for the low resolution simulation, which injects energy evenly into eight neighboring cells). Because of the slight change in injection volume, the total injected luminosity will also be modified; since we have an excellent prediction of what a simulation with a slightly different total luminosity would be (see Fig. A.5), we can control for the difference in total luminosity and isolate the influence of the injection region's shape. We compare two simulations, both with an injection radius of $0.07 H$, conductivity $\chi = 6.1 \times 10^{-3} H^2 \Omega$, and injected luminosity per volume $\ell_v = 1.42 P_0 \Omega$ but one with fiducial resolution and the other with half the fiducial resolution, resulting in total injected luminosity $2.5 \times 10^{-4} P_0 \Omega H^3$ and $1.5 \times 10^{-4} P_0 \Omega H^3$, respectively. We find that the net force per unit mass agrees between these runs at approximately the 10% level. (Note that for this value of the conductivity neither the high nor the low resolution simulation recover the analytic prediction to high accuracy.)

From Fig. A.3 it is apparent that the heating torque arises from within approximately $0.5H$ of the body for simulation L1K1, well within the size of our fiducial simulation domain. Nonetheless, the use of periodic boundary conditions does introduce artefacts that are visible in the plots of the density and pressure perturbations as structures close to the edges of the box that re-appear on the opposite side of the box (sheared, in the case of the y-edges; see Fig. A.2). To test for possible errors introduced by the use of periodic boundaries, we compared simulations in which the box size was increased to twice that of the fiducial simulation L1K1's in each direction, while maintaining all other variables constant. The measured accelerations agreed to better than 1% at every point in time, including the time (between 0.4 and 0.7 orbits)

when density perturbations had re-entered the fiducial simulation domain but had not yet reached the edge of the doubled simulation domain, as well as the steady state at later times. The same holds when the box size of L1K10 is doubled. Somewhat larger changes, at the 5% level, occur if we compare against a box with half the resolution, but three times the box size, of the fiducial simulation L1K1, likely due to the aforementioned increase in the injection region. The lack of impact that the artificially heated gas has on the measurement of the azimuthal acceleration is because the density perturbations near the edges of the box are on the order of two orders of magnitude lower than the regions closest to the body (Fig. A.2). The combined effects of the low magnitude of the density perturbations and their larger distance from the body, which reduces their contribution to the net azimuthal force, suggest that the density and pressure perturbations that exit and re-enter the box through artificial periodic boundary conditions do not impact the force calculation at the level of accuracy we are interested in here. For the extremal simulation L1K50, however, the perturbations are of the same magnitude close to and far away from the body, resulting in the increased deviation from the linear prediction seen in Fig. A.6.

Finally, we note that the simulations assume that the luminous body’s location remains fixed over a small multiple of the local dynamical timescale. In principle, for sufficiently high luminosities and local disk surface densities, the resulting torque might be able to migrate the body fast enough to invalidate this assumption. Analogous physics has been studied in the context of gravitational torques, where motion of the gravitating body can lead to a dynamical co-rotation torque and “Type III” migration [216, 217]. We do not explore this possibility here, but note that caution and additional study would be needed in any circumstance where the implied migration speed due to the heating torque exceeded a fraction of $H\Omega$.

A.4 Discussion

A.4.1 Limits of the Linear Theory

The analytic theory for the heating torque relies both on linearity, and on satisfying hierarchical separation between the scales of the displacement from co-rotation, the induced density perturbation, and the disk scale height. By numerically solving the full set of hydrodynamic equations, we can test the limits

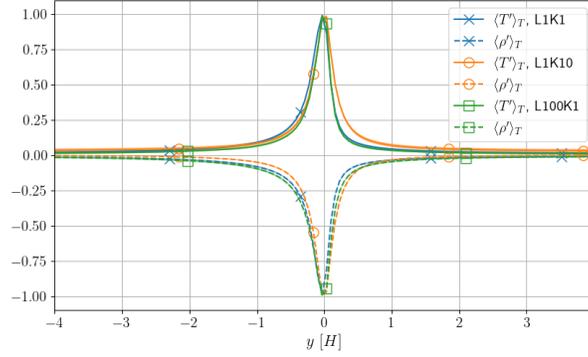


Figure A.7: Time-averaged azimuthal profiles at the body’s position of the density (dashed line) and temperature (solid line) perturbations ρ' and T' for the fiducial simulation L1K1 (blue crosses), and a high conductivity simulation L1K10 (orange dots), and a high luminosity simulation L100K1 (green squares). Each line has been normalized to its maximum value (0.14 and 0.15 times L1K1’s density and temperature perturbation maxima for L1K10; 81.3 and 137.6 times L1K1’s density and temperature perturbation maxima for L100K1) and time-averaged over the last seven orbits, $t = 3$ to 9.6 orbits. Filled-in portions denote one standard deviation over time.

of these various assumptions.

A.4.1.1 Testing the Hierarchy Requirements

The first set of assumptions are the hierarchies given by Eq. A.2 and Eq. A.3, i.e. that $x_p \ll \lambda_c \ll H$. Fig. A.6 shows how the derived acceleration scales as these assumptions are broken. The simulations closest to the analytic prediction do not have equal ratios of x_p/λ_c and λ_c/H ; rather, they prefer a smaller x_p/λ_c . As thermal conductivity increases, λ_c becomes larger whereas the offset from corotation x_p and disk scale height H stay constant. This results in a decrease in the ratio x_p/λ_c and an increase in λ_c/H , i.e. Eq. A.2 becoming better satisfied and Eq. A.3 becoming less satisfied. The result is that the characteristic wavelength of the perturbations is less well-contained by the simulation domain, particularly in the case of L1K50, whose characteristic wavelength of $3.6 H$ does not fit in the simulation domain at all. Extending L1K50’s simulation domain to a scale height where all three scales are separated by a factor of 10 (i.e. the offset from co-rotation is a factor of 100 smaller than the scale height) would require upwards of 1200 cells in each direction to resolve a single scale height, which is not even large enough to capture the full decay of higher conductivity simulations. The computational cost of such simulations is beyond the scope of this work.

A.4.1.2 Non-linear Effects

Another assumption is that the perturbations are small compared to their equilibrium values: $\rho' \ll \rho_0$. By increasing the luminosity, we can study how the acceleration departs from the linear prediction as we enter the non-linear regime where $\rho' \sim \rho_0$. The scaling relation of acceleration with fixed conductivity (Fig. A.5) shows that higher luminosity simulations that are in the non-linear regime (indicated by green squares) measure a smaller acceleration than the linear theory would predict. Because the role of periodic boundary conditions was determined to be negligible in §A.3.5, this difference could be due to either the computational issues with resolving hierarchies listed in the previous section or to non-linear effects that are not adequately captured by the linear theory. A term is considered non-linear if it contains the product of two or more perturbations, e.g. $\rho'T'$, and is hence neglected in the linear theory, which only carries the perturbations to first order under the assumption that the perturbation is much smaller than its equilibrium value. The interaction of the linear perturbations ρ' and T' through terms such as the heat flux (described below) provide a mechanism through which the simulations presented (which solve the full set of hydrodynamic equations) can differ from the linearized model of Ref. [187]. In this section we show that the measurement of the difference in azimuthal acceleration is physical, i.e. that the linear theory's neglect of higher-order terms leads to an overprediction (underprediction) of the actual non-linear net azimuthal acceleration for L100K1 (L1K10).

There are two properties of the steady-state perturbations that could contribute to the final density distribution: their profiles' shapes and their amplitude. Time-averaged azimuthal profiles of the density perturbation (Fig. A.7) in which the high luminosity (L100K1) and fiducial (L1K1) simulations' profiles lay directly on top of each other once normalized demonstrate that the shape for these two simulations is not contributing to the measured differences. Similarly, Fig. A.8 shows the non-linear term $\nabla \cdot [\chi\rho'\nabla T']$ (the divergence of the heat flux's contributions by non-linear terms; this term is subtracted from the time derivative of the internal energy density in Eq. A.11). The normalized profiles of the heat flux from perturbed quantities of both the fiducial (L1K1) and high luminosity (L100K1) simulations overlay one another within one standard deviation. These similarities in shape show that the relevant length scales are indeed the same

for these two simulations, further suggesting that it is not the shape but rather the amplitude of the profiles that contributes to the measured differences. Indeed, the magnitude of the higher luminosity run's heat flux is four orders of magnitude larger than the fiducial simulation. In contrast, the azimuthal density and heat flux profiles of the high conductivity run L1K10 (which has a larger characteristic wavelength $\lambda_c \propto \chi^{1/2}$) do not line up with the fiducial and high luminosity runs even when normalized. In the high conductivity run, the heat more readily diffuses from its injection region to the $y > 0$ region, resulting in a broader high temperature region with a lower maximum temperature compared to the L1K1 and L100K1 simulations (Fig. A.7). For the same reason of heat diffusion, the azimuthal profile of density perturbations in Fig. A.7 appears more symmetric close to the body than L1K1 and L100K1. As discussed below, the increased symmetry results in a smaller net acceleration close to the body in L1K10 than in L1K1. This change in the shape of the density's azimuthal profile is not captured by the linear theory, which predicts that the three-dimensional distribution of the density perturbation does not depend on thermal conductivity [187, Eq. 119].

To quantify what the shape of the heat flux and density/temperature azimuthal profiles means for the acceleration of the body, Fig. A.9 plots the net azimuthal acceleration from the gas as a function of distance from the body (very similar to Fig. A.3's plot of the sum of the azimuthal acceleration in either y -direction), normalized to the values of the fiducial run at every point. Because the linear theory predicts that only the size of the low density perturbation and not the shape should change with varying luminosity and conductivity, we use the fiducial simulation as a template to predict the azimuthal acceleration profile and scale it by L/χ . These profiles are plotted as horizontal lines in Fig. A.9. In this figure, the high luminosity (L100K1) simulation's acceleration profile is modified only close to the body, where it is lower than might be expected from linear theory. Notably, the magnitude of the perturbations is also the highest close to the body (as seen in Fig. A.2), suggesting that non-linear effects such as the heat flux term profiled in Fig. A.8 cause the deviation from linear theory. On the other hand, the azimuthal acceleration on the body in L1K10 is less than the linear prediction within about half a scale height of the body as suggested by the symmetry seen in Fig. A.7, but grows steadily due to higher value of the density perturbations at $y > 0.5 H$ compared to those at $y < -0.5 H$. The azimuthal acceleration saturates at a value approximately

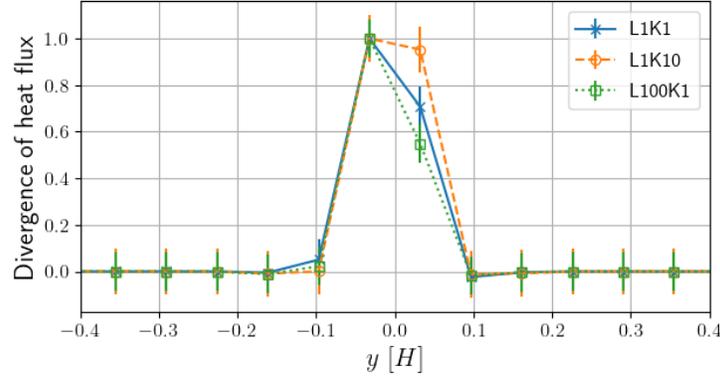


Figure A.8: Time-averaged azimuthal profiles at the body’s position of the divergence of the heat flux’s non-linear contribution $\nabla \cdot [\chi \rho' \nabla T']$ for the fiducial simulation L1K1 ($\ell_v = 1.42 P_0 \Omega$, $\chi = 0.017 H^2 \Omega$; blue crosses), the high conductivity simulation L1K10 ($\ell_v = 1.42 P_0 \Omega$, $\chi = 0.17 H^2 \Omega$; orange circles), and the high luminosity simulation L100K1 ($\ell_v = 142 P_0 \Omega$, $\chi = 0.017 H^2 \Omega$; green squares). Each line has been normalized to its maximum value (0.004 and 2.9×10^4 times the fiducial simulation’s maximum value, respectively) and time-averaged over $t = 3$ to 9.6 orbits. Error bars denote one standard deviation over time.

twice as high as the linear theory, ultimately leading to a measured azimuthal acceleration larger than the linear prediction (see Fig. A.6). Although the increased value of the acceleration far from the body might suggest that periodic boundary conditions are artificially increasing the measured acceleration, this density distribution was achieved in larger boxes before the density perturbations re-entered the simulation domain and is thus physical (see §A.3.5), suggesting that it is indeed physics neglected by the linear theory that alter the symmetry of the density distribution.

We again note that the separation of the effect of the body’s luminosity from its gravitational potential is only valid in the linear regime. In the non-linear regime, interaction between these two effects (which linearly act in the same way to provide a net outward migration) could result in deviation from the linear prediction. Exploring aspects of this interaction is left to future studies.

A.4.2 Physical Parameter Regimes

A.4.2.1 Stars or Accreting Compact Objects in a Thin Disk

Migration processes may be important in geometrically thin AGN accretion disks. Stars may form within such disks as a consequence of gravitational instability [201, 202], and they may also be captured

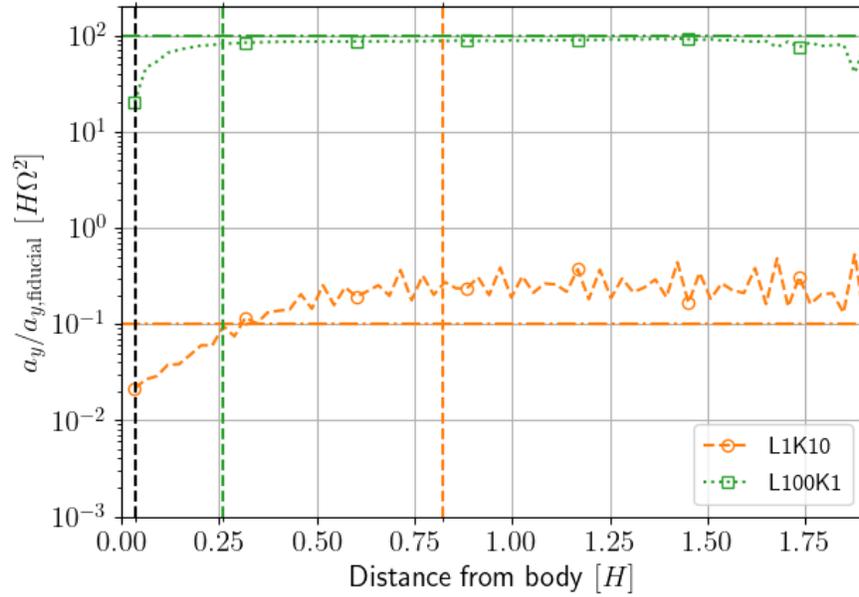


Figure A.9: Snapshot of the radial profile of the net azimuthal acceleration on the body at a time of 9 orbits for a high conductivity simulation L1K10 ($\ell_v = 1.42 P_0\Omega$, $\chi = 0.17 H^2\Omega$; orange dash-dot line), and a high luminosity simulation L100K1 ($\ell_v = 142 P_0\Omega$, $\chi = 0.017 H^2\Omega$; green dotted line), normalized at every point to the fiducial simulation L1K1. As in Fig. A.3, the injection radius $r_{\text{rad}} = 0.04 H$ for all three simulations (black vertical line) and half the characteristic wavelength for L1K1 and L100K1 ($\lambda_c/2 = 0.26 H$) and L1K10 ($\lambda_c/2 = 0.82 H$) are plotted. Horizontal lines are the profiles that L1K10 and L100K1 would have if they had the same shape as L1K1, i.e. the linear theory's prediction.

from a cluster whose orbits intersect the disk gas [199]. Either circumstance could lead to a population of luminous stars, or accreting stellar-mass compact objects, orbiting within an AGN disk.

The full ramifications of having a population of stellar-mass objects within AGN disks are complex, and we do not discuss them here. Rather, we assume that we have a single luminous object orbiting on a circular, non-inclined orbit, with the same sense of rotation as the disk gas. The question we seek to answer is whether the heating torque is large enough, compared to previously studied torques arising at the Lindblad and co-orbital resonances, that it should be included in models of migration within AGN disks.

We assume, consistent with our numerical results, that the analytic result given as Eq. A.7 provides a good estimate of the ratio of the heating torque, Γ_{heat} , to the fiducial torque scaling, Γ_0 . We assume a Keplerian disk ($q = 3/2$) in which the pressure $p_0 \propto r^{-n}$. At $r = r_0$ the relevant quantities can then be written as,

$$x_p = \frac{nc_s^2}{3r_0\Omega_0^2\gamma}, \quad (\text{A.16})$$

$$\lambda_c = 2\pi\sqrt{\frac{2\chi}{3\gamma\Omega_0}}, \quad (\text{A.17})$$

$$L_c = \frac{4\pi}{\gamma}GM\chi\rho_0. \quad (\text{A.18})$$

Here χ is the thermal diffusivity in the disk gas surrounding the luminous object. If the diffusivity is physically the result of radiative diffusion, we can write [192, correcting the factor of 4 typo],

$$\chi = \frac{16\gamma(\gamma - 1)\sigma T^4}{3\kappa\rho_0^2H^2\Omega_0^2}, \quad (\text{A.19})$$

where κ is the opacity, T is the temperature, and σ is the Stefan-Boltzmann constant.

For both massive stars and accreting compact objects, the Eddington luminosity provides a very rough but reasonable estimate of how the likely luminosity scales with the mass. We write,

$$L = \epsilon L_{\text{Edd}} = \frac{4\pi\epsilon cGM}{\kappa}, \quad (\text{A.20})$$

where ϵ is an efficiency factor that may be larger than one. The above formulae then give a scaling,

$$\frac{\Gamma_{\text{heat}}}{\Gamma_0} \propto \left(\frac{H}{r_0}\right) c_s^2 \kappa^{1/2} \frac{H^2 \rho_0^2 \Omega_0^{3/2}}{T^6}. \quad (\text{A.21})$$

The numerical pre-factor depends upon the assumed vertical structure of the disk. Taking $\rho_0 = \Sigma/H$ and $c_s = \gamma H^2 \Omega_0^2$ we find,

$$\frac{\Gamma_{\text{heat}}}{\Gamma_0} \simeq 0.053 \frac{n\epsilon}{\gamma(\gamma-1)^{1/2}} \left(\frac{H}{r_0}\right) \frac{c\kappa^{1/2}}{\sigma^{3/2}} \frac{H^2 \Sigma^2 \Omega_0^{7/2}}{T^6}. \quad (\text{A.22})$$

There is no dependence on the mass of the luminous object. We note that the above analysis has assumed in various places that the disk is optically thick, that it is supported by gas pressure, and that the luminosity of the embedded object is transported out by radiative diffusion.

Given a disk model, for example a Shakura-Sunyaev disk in one of the gas pressure dominated regimes [9], it is straightforward to estimate the ratio of $\Gamma_{\text{heat}}/\Gamma_0$. The result is fairly complex expressions that obscure the basic question of whether Γ_{heat} can be neglected when considering migration. It is more illuminating to forego explicit reference to the opacity, and write Eq. (A.22) in a manifestly dimensionless form that involves the optical depth τ . To do so we need the following results for a thin accretion disk in a steady state [46]:

$$\tau = \Sigma\kappa, \quad (\text{A.23})$$

$$T^4 = \frac{3\tau\dot{M}\Omega_0^2}{8\pi\sigma}, \quad (\text{A.24})$$

$$\nu\Sigma = \frac{\dot{M}}{3\pi}, \quad (\text{A.25})$$

$$\nu = \alpha c_s H. \quad (\text{A.26})$$

(We have dropped some unimportant numerical factors from these expressions.) Using these results, and adopting reasonable values for the adiabatic index and pressure gradient parameter ($\gamma = 5/3$, $n = 3$) we obtain,

$$\frac{\Gamma_{\text{heat}}}{\Gamma_0} \sim 0.07 \left(\frac{c}{v_K}\right) \epsilon \tau^{-1} \alpha^{-3/2}, \quad (\text{A.27})$$

where v_K is the Keplerian velocity in the disk. The condition for the heating torque to be important (relative to the Lindblad and co-orbital torques, again as defined by [187] rather than [192]), $\Gamma_{\text{heat}} > \Gamma_0$) is then,

$$v_K \tau \alpha^{3/2} \lesssim 0.07 c \epsilon. \quad (\text{A.28})$$

In an AGN disk we expect $\alpha < 1$, and across most of the region where stars would form or be captured $v_K \ll c$. It is then clear that an embedded object, radiating a luminosity of the order of the Eddington limit ($\epsilon \sim 1$), will experience dominant heating torques at any radii where the optical depth is modest.

As a specific example, we apply the criterion in Eq. A.28 to the model of a constant Toomre parameter disk [201]. Taking the standard parameters presented in this model, we use a central black hole mass of 10^8 solar masses, a radiative efficiency $\mathcal{E} = L/\dot{M}c^2 = 0.1$, assume that the viscosity is proportional to the total pressure rather than the gas pressure, and use the electron-scattering opacity $\kappa = 0.4 \text{ cm}^2/\text{g}$. Although this model cannot be directly mapped onto a Shakura-Sunyaev disk, it is useful as an example. From Ref. [201]’s Eq. 10, the radius at which gravitational instability sets in is at $r_{Q=1} \approx 10^3 R_s$. Using a value of $\alpha = 0.01$ appropriate for MRI turbulence and assuming a luminosity on the order of the Eddington limit, the resulting bound is

$$\tau \lesssim 2200. \quad (\text{A.29})$$

Using Ref. [201]’s Eqns. 16 and 18 to calculate the temperature and surface density at $r_{Q=1}$ and the prescription in Eq. A.19 for thermal conductivity, the relevant length scales are

$$x_p \sim 0.02 R_s \quad (\text{A.30})$$

$$\lambda_c \sim 8 R_s \quad (\text{A.31})$$

$$H \sim 10 R_s. \quad (\text{A.32})$$

The assumption that $x_p \ll \lambda_c < H$ holds for this model of a luminous object in an AGN disk, although the characteristic wavelength λ_c and the gas scale height H are the same order of magnitude, similar to the simulations presented in this work (see Table A.3).

A.4.2.2 Low-mass Planets in a Protoplanetary Disk

Thermal torques were originally proposed and studied in the context of low-mass planets in protoplanetary disks, where under some circumstances they can be of the same order of magnitude as Lindblad torques. Prior simulations focused on this regime include the work of Ref. [185] and that of Ref. [186]. There are physical differences between these simulations and those presented in this paper. In particular, in Ref. [185] the planet has mass but no luminosity (and hence is affected by only the cold thermal torque), while in Ref. [186] the planet is both luminous and massive (and hence is affected by both the cold thermal and the heating torque). Table A.3 compares parameters for these two simulations and the simulation L1K1.

	Ref. [185]	Ref. [186]	Fiducial Simulation (L1K1)
λ_c	2 cells [†] /0.014 AU [†]	2.34 cells [§] /0.0238 AU [†]	41.3 cells/0.084 AU
x_p	0.85 cells/0.006 AU [†]	0.98 cells [§] /0.01 AU [*]	6 cells/0.0122 AU
H	28 cells/0.196 AU	21.3 cells [§] /0.19 AU [*]	62 cells/0.126 AU
L [erg/s]	N/A	$6.0 \times 10^{27\dagger}$	6.0×10^{30}
χ [cm ² /s]	1.5×10^{15} cm ² /s [†]	$4.35 \times 10^{15\dagger}$	1.0×10^{17}
λ_c/x_p	2.33	2.38	6.88
H/λ_c	14 [†]	8	1.5

Table A.3: Summary of important values for different simulations. Ref. [185] studies the cold thermal torque (no luminosity), while Ref. [186] studies the cold and heating torques in a semi-global simulation. The values in the table were obtained either from Sec. 5.3.2 in [187] (†), by private communication (*; in particular $H/R = 0.036$), or by calculation from parameters found in the respective work (§). Unmarked values are straightforward calculations from other values. Physical values for the fiducial simulation L1K1 were calculated assuming a body orbiting at 5.2 AU (the same as [186]) around a solar-mass central body. Here $\lambda_c \equiv k_c^{-1}$ rather than as in Eq. A.4 in accordance with Ref. [187]’s definition.

Although Ref. [185] and Ref. [186] are more comprehensive physically than the present study, it is useful to compare the results of these two works to those presented in §A.3 to investigate the importance of resolving the region within one characteristic wavelength of the body. Additional studies investigate global effects such as the excitation of orbital eccentricities and inclinations [189, 194, 212] and horseshoe streamlines [213], which the local unstratified simulations in the present study do not capture; for instance the gas flow that would be significantly altered around a hot planet [213] is not seen in the streamlines of L100K1 plotted in Fig. A.2. Though these global-scale simulations could be subject to the same limits of resolution as Ref. [185] and Ref. [186], they are not discussed in detail for the sake of clarity and brevity.

Although the results of both Ref. [185] and Ref. [186] highlighted the importance of thermal effects for an accurate assessment of the migration rate, there was a significant mismatch between the quantitative values obtained and the subsequent analytic theory of Ref. [187]. Inspection of Table A.3 and the results of §A.3.5 suggest that this discrepancy may well be due to the difficulties inherent in resolving the relevant scales in a global simulation. Compared to previous simulations [185, 186], our fiducial simulation L1K1 better resolves the characteristic wavelength of the density perturbation. Ref. [185] has approximately two cells spanning λ_c [187], whereas Ref. [186] has approximately 2.3 cells to resolve λ_c . By using a local domain, our fiducial simulation L1K1, which resolves λ_c with 41 cells, is better poised to capture the full effect of

the thermal torques, since as Fig. A.3 and A.9 show, most of the contribution to the azimuthal acceleration comes from material within $\lambda_c/2$ of the body. Limited resolution could be one of the reasons that Ref. [186] sees a net force about an order of magnitude below the predicted value [187]. As for physical parameters, our fiducial simulation L1K1’s luminosity is three orders of magnitude larger than those presented in Ref. [186], leading to a large value for the heating torque. However, the results should continue to scale down to lower values of luminosity, where the assumption that the planet does not change its distance from corotation over the course of the simulations should be more accurate.

Our simulations also explore a different regime in terms of the scale hierarchy. The simulation of Ref. [185] has $H/\lambda_c = 14$, whereas our simulations have $H/\lambda_c = 1.5$; similarly, in Ref. [185], $\lambda_c/x_p = 2.3$ and in our fiducial simulation L1K1 $\lambda_c/x_p = 6.88$ (Table A.3). We have better scale separation between λ_c and x_p at the expense of less scale separation between λ_c and H . Our closer-to-linear results support our argument that the first criterion is more essential to the linear theory than the second. As discussed in §A.4.1.1, the restriction on scale height is not a physical requirement but rather an ease-of-computational one; therefore this work simply explores a slightly different physical parameter regime.

We conclude that poor resolution of the characteristic wavelength contributes to the mismatch between previous simulations’ measurement of the heating force and the linear theory’s prediction. This conclusion is supported by the resolution study in §A.3.5 wherein a decrease in resolution led to a decrease in the measured value of azimuthal acceleration to below the linear prediction, similar to how Ref. [186]’s measurement was an order of magnitude smaller than the predicted value [187]. Using a higher resolution (possible in a local setup), we obtain agreement to within 10% with the analytic theory. The remaining discrepancies are usually small, exceed the analytic prediction, and reduce with higher resolution. Exceptions are simulations with thermal conductivity much smaller than the fiducial value. This trend suggests that a more precise value depends on the exact physics close to the body, which would be better modelled with a self-consistent luminosity prescription to capture accretion onto the body. Compared to previous global studies, in local shearing box simulations, the relevant small scales are better resolved and the net azimuthal acceleration is within 10% of the linear theory in the linear regime.

A.5 Conclusions

In this work, we have used three-dimensional local simulations of a zero-mass test particle in an inviscid unstratified disk to test the analytic theory for the heating torque developed by Ref. [187]. The heating torque arises from the interaction between a luminous disk-embedded body and Keplerian shear, which distorts low-density regions that were heated by the body into asymmetric lobes that exert a net gravitational force on the planet. In the regime where the resulting density perturbations are linear, we find good agreement between the results of our direct numerical simulations and the analytic theory. We surmise that prior global simulations probably lacked enough resolution of the energy injection region, leading to an under-estimate of the magnitude of the heating torque. Going beyond the linear theory, we explored regimes of high thermal conductivity and high luminosity. We find that at high luminosity the derived torque is smaller than the linear prediction, and attribute this as being due to non-linear terms in the heat flux. In the high conductivity regime we infer a higher acceleration than predicted by the linear theory. We find that both the non-linear terms in the heat flux and computational limitations contribute to this larger value.

At the linear level the heating torque can be considered separately from other contributions to the migration of disk-embedded bodies. Although numerically convenient, there are few if any physical circumstances where gravitational [190] and other thermal torques [185] would not also need to be considered. In most cases, study of these torques requires a combination of analytic theory, local, and global numerical simulations, whose results can partially be encapsulated in relatively simple torque formulae [192, 218]. In the context of low-mass planet migration, using such formulae, the heating torque is estimated to be most important (relative to the sum of all other torques) for masses on the order of one Earth mass [187]. This study neglects additional physics to better isolate the impact of the heating torque on low-mass planet migration; inclusion of the body's gravitational potential, disk stratification, and viscosity among other physics can be included in later studies.

A second environment where the heating torque might be important is for the migration of luminous objects (massive stars or accreting compact objects with luminosities of the order of the Eddington limit) in AGN disks. Using simple scaling arguments, and the analytic theory of Ref. [187], we find that the

heating torque is expected to provide a dominant contribution to the total migration torque at disk radii where the optical depth drops below a critical value. The heating torque may therefore impact models for the migration, trapping, and growth of objects embedded within AGN disks, and should be considered in future analyses of such systems. In the case where the disk-embedded body is itself an accreting compact object, the mechanical luminosity of outflows may also modify the local density distribution and generate a migration torque [219].

Using local simulations on a uniform grid, we have been able to verify the analytic predictions for the strength of the heating torque at approximately (in the most favorable cases) the 10% level. More precise tests would be possible using static mesh refinement methods, which would also allow a fuller mapping of how the thermal torque scales with the control parameters in regimes where the assumptions of the analytic theory fail. It would also be valuable to relax the assumptions of a constant energy injection rate and conductive energy transport. Simulations that consistently resolve accretion onto disk-embedded objects, and the radiative feedback that accretion produces, are challenging but are becoming increasingly feasible.

A.6 Acknowledgements

We thank Pablo Benítez-Llambay and the anonymous referee for useful discussions and feedback. AMH acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1650115. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. PJA acknowledges support from NASA TCAN award 80NSSC19K0639. This research was carried out in part during the 2019 Summer School at the Center for Computational Astrophysics, Flatiron Institute. The Flatiron Institute is supported by the Simons Foundation. The numerical simulations and analysis were performed on the Rusty cluster of the Flatiron Institute.

Appendix B

The Kerr Metric and Thin Accretion Disk Properties

B.1 Introduction

In this note, I will cover some properties of the Kerr metric and Novikov-Thorne [49] thin disks that are particularly relevant to my research involving gas dynamics within the innermost stable circular orbit (ISCO) around a black hole. A given black hole in astrophysics carries two properties^a in addition to its spatial location and velocity relative to some observer: its mass M and its angular momentum J , which I often refer to via the dimensionless spin $a \equiv J/M$ ^b. I define the coordinate system such that the black hole angular momentum \vec{J} points along the z -axis.

In Boyer-Lindquist coordinates, the Kerr metric reads [220]:

$$ds^2 = -\frac{\Delta}{\rho^2} \left(dt - a \sin^2 \theta d\phi \right)^2 + \frac{\sin^2 \theta}{\rho^2} \left((r^2 + a^2) d\phi - a dt \right)^2 + \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2, \quad (\text{B.1})$$

where

$$\Delta = r^2 - 2Mr + a^2 \quad (\text{B.2})$$

$$\rho^2 = r^2 + a^2 \cos^2 \theta. \quad (\text{B.3})$$

Throughout this chapter, I will refer to three different frames:

- (i) Coordinate frame. The Boyer-Lindquist coordinate time is labelled as t .

^a Astrophysical black holes probably have charge $Q = 0$, because any charge will attract material with opposite charge over a short timescale.

^b I use geometrized units $G = c = 1$ in this section.

- (ii) Locally non-rotating frame (LNRF; [221]). An observer at Boyer-Lindquist coordinates (t, r, θ, ϕ) will have proper time labelled as τ . This frame includes the effects of frame-dragging. It is also known as the Eulerian frame or the zero-angular momentum observer (ZAMO).
- (iii) Fluid frame. This frame only exists in the presence of fluid comprising an accretion disk and depends on the 4-velocity of the fluid at Boyer-Lindquist coordinates (t, r, θ, ϕ) . It will be labelled with proper time τ_0 .

B.1.1 Properties of the Kerr Metric

B.1.1.1 Prograde, Midplane Surfaces in the Kerr Metric

For simplicity, in this section I will only consider surfaces in the midplane of the black hole relevant for particles with angular momentum aligned with the black hole's spin (i.e. no retrograde orbits).

The first important surface is of course the (outer) event horizon, which occurs as a coordinate singularity in Boyer-Lindquist coordinates when the metric component g_{rr} goes to infinity. From Eq. B.1, we see that this singularity occurs when $\Delta \rightarrow 0$, i.e. $0 = r_H^2 - 2Mr_H + a^2$. Taking the positive root for the outer event horizon, we have

$$r_H = M + \sqrt{M^2 - a^2}. \quad (\text{B.4})$$

Any light or massive object that enters the event horizon cannot escape from it.

The next surface, the ergosphere, defines where the black hole drags spacetime along with it at the speed of light, such that a particle must move in the direction of black hole spin. Because this forced rotation means that particles gain energy, particles that escape the ergosphere can effectively remove rotational energy from the black hole itself (known as the Penrose process [222]). Extracting rotational energy from an astrophysical black hole can power relativistic jets [223] and limit the spin-up of black holes [224]. The ergosphere sits outside the event horizon at radius r_E where [220]:

$$(r_E - M)^2 = M^2 - J^2 \cos^2 \theta. \quad (\text{B.5})$$

At the midplane, $\theta = \pi/2$ gives

$$r_E = 2M. \quad (\text{B.6})$$

Interestingly, the ergosphere location is independent of black hole spin.

Much of my research, many accretion disk models, and many methods for measuring black hole spin rely on the location of the innermost stable circular orbit (ISCO). The ISCO is defined as the smallest radius where a massive test particle can still have a stable circular orbit around the black hole. For $r < r_{\text{ISCO}}$, the test particle's orbit becomes predominantly radial rather than circular, earning the region within the ISCO the name “the plunging region”. The ISCO lies at radius [220]:

$$r_{\text{ISCO}} = 3 + Z_2 - \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)} \quad (\text{B.7})$$

where

$$Z_1 = 1 + (1 - \chi^2)^{1/3} \left((1 + \chi)^{1/3} + (1 - \chi)^{1/3} \right) \quad (\text{B.8})$$

$$Z_2 = \sqrt{3\chi^2 + Z_1^2} \quad (\text{B.9})$$

and $\chi = J/M^2 = a/M$.

A couple more surfaces I will discuss here also rely on the concept of orbits around black holes. In particular, at radii beyond the marginally bound radius r_{mb} , a massive particle can still have a bound orbit; it just will not be a stable circular orbit. Finally, the photon sphere determines the last location r_{ph} of a circular orbit for a massless particle (i.e. a photon). The photon sphere was particularly important in measuring the black hole shadow [225]. These surfaces are given by [220]:

$$r_{mb} = (1 + \sqrt{1 + \chi})^2 \quad (\text{B.10})$$

$$r_{ph} = 2 \left[1 + \cos \left(\frac{2}{3} \cos^{-1} \chi \right) \right]. \quad (\text{B.11})$$

Fig. B.1 illustrates the location of these surfaces as a function of black hole spin. For a low spin black hole ($a \lesssim 0.94$), a portion of the plunging region (volume within the ISCO) lies outside the ergosphere, i.e. there are no stable circular orbits for massive particles that experience frame dragging; they are all unstable. For $a \gtrsim 0.94$, the entire ISCO is within the ergosphere, i.e. there are stable circular orbits that experience frame dragging.

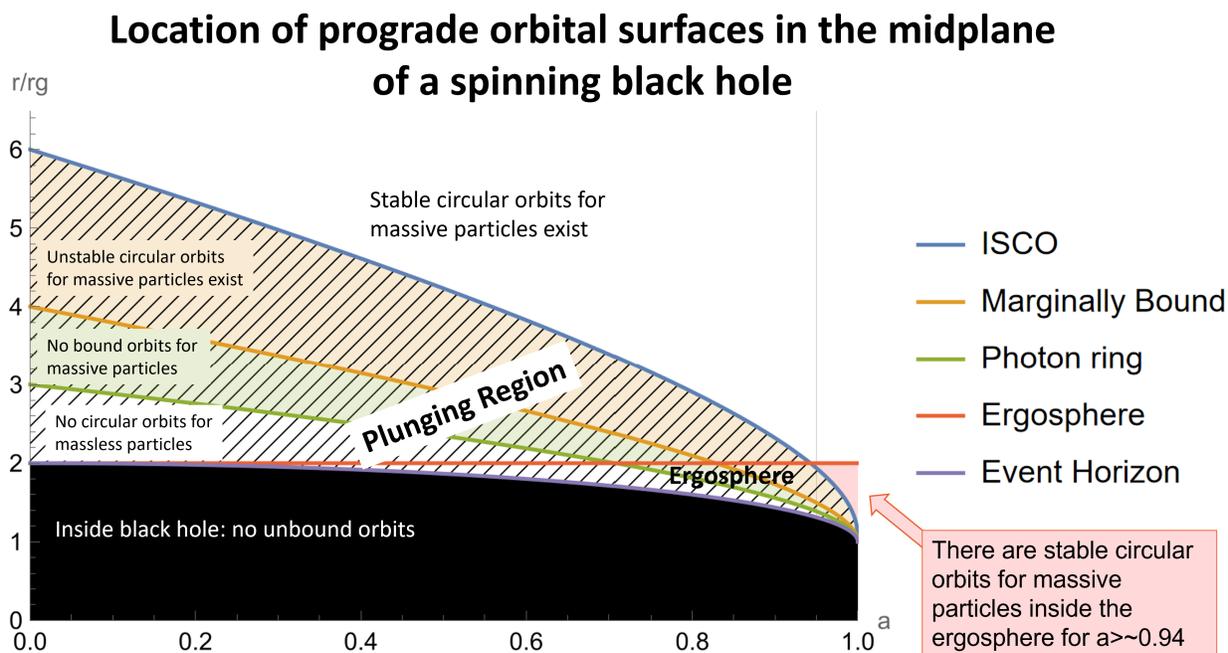


Figure B.1: Black hole surfaces as a function of black hole spin a . The hatched area with $r < r_{ISCO}$ is the plunging region (only unstable orbits). The yellow-filled region with $r_{mb} < r < r_{ISCO}$ hosts unstable circular orbits, while the green-filled region with $r_{ph} < r < r_{mb}$ has unstable bound orbits. At $r < r_{ph}$, no circular orbits are possible, even for massless particles such as photons. The red region shows where there are stable circular orbits within the ergosphere ($r_{ISCO} < r_E$), which only occurs for large spins $a \gtrsim 0.94$.

B.1.1.2 The Lapse Function

The lapse function α is a property of the metric itself and is independent of the accretion disk flow. It describes how much spacetime is bent: far from the black hole, the lapse function approaches 1, whereas close to the event horizon it approaches 0. The lapse function is defined as:

$$\alpha \equiv \frac{d\tau}{dt} = \frac{1}{\sqrt{-g^{tt}}}. \quad (\text{B.12})$$

Fig. B.2 shows the lapse function over radius for two different values of black hole spin.

B.1.2 Relativistic Thin Disk Properties

In this section, I will outline some important properties of the relativistic thin disk model [49]. This model makes a variety of assumptions, including local thermodynamic equilibrium such that local dissipation immediately radiates away.

In curved spacetime, the fluid Lorentz factor is not simply the zeroth component of the 4-velocity. In fact, the Lorentz factor is defined relative to the LNRF as:

$$\gamma \equiv \frac{d\tau}{d\tau_0}. \quad (\text{B.13})$$

Meanwhile, the 4-velocity is given by

$$u^\mu \equiv \frac{dx^\mu}{d\tau_0}, \quad (\text{B.14})$$

and in particular that means the zeroth component reads

$$u^0 = \frac{dt}{d\tau_0} = \frac{dt}{d\tau} \frac{d\tau}{d\tau_0} \quad (\text{B.15})$$

$$= \frac{1}{\alpha}, \quad (\text{B.16})$$

where the lapse function is defined in Eq. B.12. See Ref. [226, Eq. 5.28]. The Lorentz factor for the Novikov-Thorne model is shown in Fig. B.3a (see their Eq. 5.4.4b). A nonspinning black hole reaches Lorentz factors of 1.15 at the ISCO ($r = 6r_g$), whereas a black hole with $a = 0.95$ reaches Lorentz factors close to 1.3 at the ISCO ($r \sim 2r_g$).

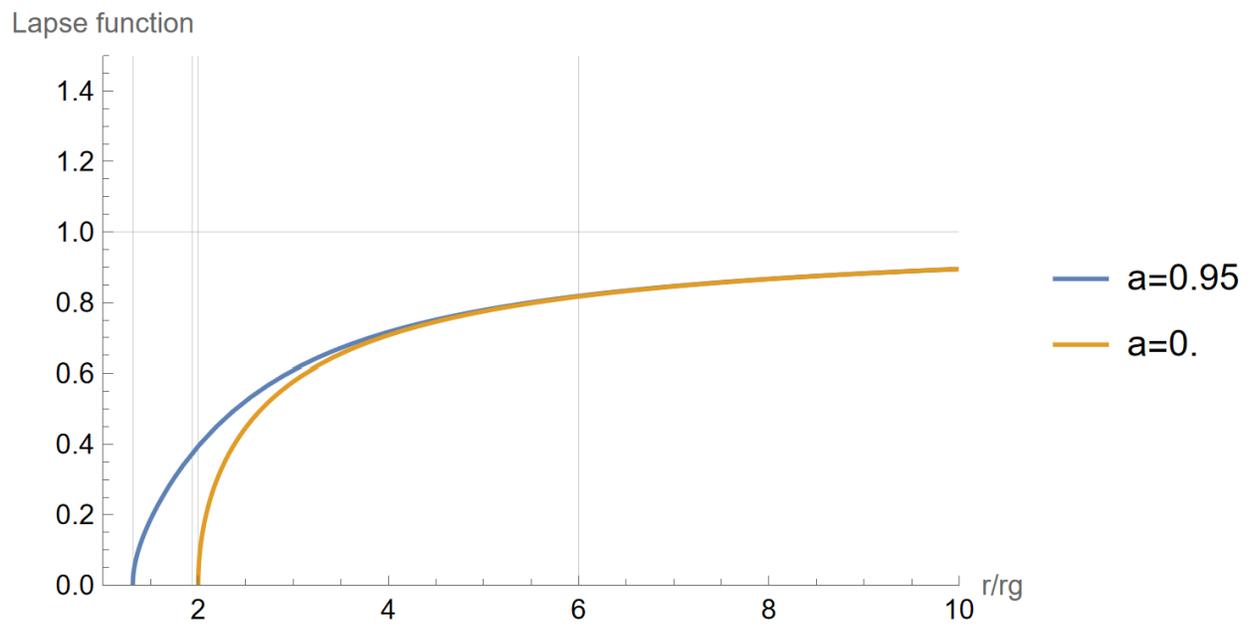


Figure B.2: Lapse function (Eq. B.12) over radius for two different black hole spins a .

Imagine a particle with a stationary orbit such that it has constant (r, θ) and a uniform angular velocity

$$\Omega = \frac{d\phi}{dt}. \quad (\text{B.17})$$

Note that this stationary frame does not have to coincide with the LNRF frame, though it can for a certain value of Ω . A distant (asymptotic) observer sees the particle complete one orbit in a time $t_{\text{orbit}} = 2\pi/\Omega$; in this sense, Ω is the angular velocity measured from infinity [220, Problem 33.2]. The particle (or a stationary observer) measures the time to complete an orbit as

$$\tau_{\text{orbit}} = t_{\text{orbit}} \left(-g_{tt} - 2\Omega g_{t\phi} - \Omega^2 g_{\phi\phi} \right)^{1/2}. \quad (\text{B.18})$$

For an LNRF observer with $\Omega = -g_{\phi t}/g_{\phi\phi}$, the factor in parentheses is simply the lapse function defined in Eq. B.12.

The gas at each radius orbits with the Keplerian angular frequency Ω_K , modified by a relativistic correction [220, Eq. 5.4.3]:

$$\Omega_K = \frac{M^{1/2}}{r^{3/2} + aM^{1/2}}. \quad (\text{B.19})$$

Fig. B.3b shows the ratio $\tau_{\text{orbit}}/t_{\text{orbit}}$ for a stationary observer with angular velocity Ω_K as a function of distance from the event horizon for a nonspinning and rapidly-spinning black hole. The measured time over an orbital period by an asymptotic observer is about a factor of three greater than the proper time for a stationary observer at the ISCO of a black hole with spin $a = 0.95$, and about 1.4 times greater for a stationary observer at the ISCO of a nonspinning black hole.

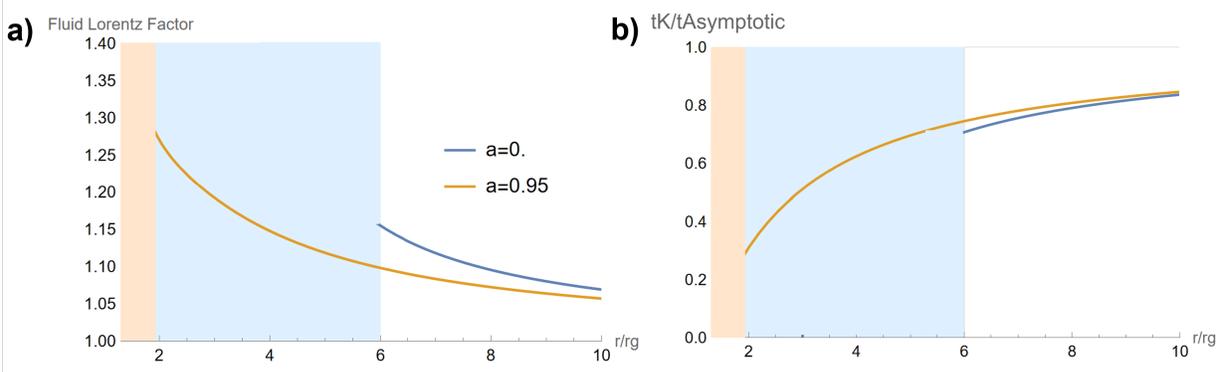


Figure B.3: a) Lorentz factor γ for the Novikov-Thorne relativistic thin disk as a function of radius for two dimensionless spins. b) Ratio of fluid proper time to observer at infinity for gas with angular velocity given by the Novikov-Thorne thin disk model.