ON ELECTROMAGNETIC OBSERVABLES FROM SUPERMASSIVE BLACK HOLE ACCRETION FLOWS

BY

GEORGE NATHANIEL WONG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois Urbana–Champaign, 2021

Urbana, Illinois

Doctoral Committee:

Professor Gilbert Holder, Chair
Professor Charles F. Gammie, Director of Research
Professor S. Lance Cooper
Associate Professor Xin Liu

Abstract

Relativistic jets have been observed to originate from the centers of many galaxies. It is likely that the jets are powered by spinning supermassive black holes via a dynamical interaction between magnetic fields close to the hole and the warped spacetime predicted by general relativity. This dissertation describes a series of projects aimed at understanding and identifying signatures of the physical quantities relevant to the black hole–jet connection in both observational and theoretical contexts.

I start with a review of astrophysical black hole accretion systems and the radiative physics that governs the generation of electromagnetic signals from hot leptons near the hole. I then describe the numerical tools I use to simulate the accretion and generate synthetic images and spectra, paying particular attention to my contributions and extensions to the code. Next, I discuss my contribution to the theoretical analysis of the first event-horizon-scale black hole accretion flow images, which were produced by the Event Horizon Telescope.

The remainder of the dissertation covers projects designed to support a theory-based guide for the next generation of electromagnetic black hole observation in the context of the jet–hole connection. I begin by describing two projects focused on understanding the composition of the jet near the hole. The first project studies mass entrainment through the jet–disk boundary layer as a mechanism to feed the jet at small scales. The second project studies electron–positron drizzle pair creation due to the background radiation field produced by the hot accretion flow. I conclude with a discussion of black hole glimmer, a novel universal signature of black hole spin that can be measured from high-resolution black hole movies and used to determine the orientation and magnitude of a black hole's angular momentum vector.

Acknowledgments

Having Charles Gammie as my advisor has been my greatest fortune as a graduate student. Charles is a seemingly inexhaustible fount of knowledge and wisdom, and he has been the epitomical scientific role model. Each of my projects was buttressed by his deep understanding of and intuition for the astrophysical and the computational, and his unwavering support and confidence are to thank for my growth as a researcher. I can only hope to one day be as effective a mentor as he.

Next, I would like to thank Josh Dolence for hosting me at Los Alamos National Lab midway through my graduate career and giving me a semester that was educative, productive, and incredibly enjoyable. I owe Ben Ryan for helping to establish the foundations of my magnetohydrodynamical black hole knowledge at Illinois. Thanks also to Jonah Miller and my office mates at Los Alamos, Brooks Kinch and Greg Salvesen, who taught me a lot about science and postdoctoral life.

I am indebted to Michael Johnson for both his invitation to visit the Black Hole Initiative and his continuing support. The Harvard trip launched collaborations with many amazing scientists, all of whom played an important role in helping me develop into an effective scientist. I especially note Daniel Palumbo, Alex Lupsasca, Shahar Hadar, and Ramesh Narayan, with whom I have had hours of scientific and vocational discussions.

Membership in the Event Horizon Telescope and Horizon collaborations has been crucial on my academic path, and I thank the many members of each collaboration for the untold days of telecons and stimulating discussion. I look forward to continuing the collaboration in person during my time at the Institute for Advanced Study and the Princeton Gravity Initiative, and I especially thank Jim Stone and Frans Pretorius for the kind and compelling personal invitations that accompanied their formal job offers.

I am deeply grateful to Ben Prather, who was my original compatriot in the EHT and who has attended with distinction the Sisyphean task of tolerating my garrulous, nonsensical loquacity. I thank my academic twin, Patrick Mullen, for teaching me about the Moon and for being an exemplary scientific sounding board. I am excited for our future collaboration at the IAS. Thank you to Joshua Yao-Yu Lin for repeatedly thinking of and inviting me to get lunch and discuss new projects, especially when we were both new graduate students. Thank you also to Vedant Dhruv for meticulously checking everything I do, Eric Petersen for extensive historical and political discussions, and the other members of the group, Abhishek Joshi, David Lee, and Allycia Gariepy, all for their tolerance of my tangents during group meetings and for humoring me when I would show up unannounced to their desks. I am especially thankful for the opportunity to work with some of the amazing non-doctoral-student members of the Gammie group, Ricardo Yarza, Yufeng Du, and Alice Maurel. I express particular gratitude to Alice for the insightful questions, discussions, and friendship that helped me stay grounded in the demanding months preceding the EHT results and after.

The penultimate calendar year of my graduate studies was disturbed by the COVID-19 pandemic, and my astrophysical research was halted as I worked to understand the virus spread with a group of wonderful people and scientists—including Nigel Goldenfeld, Sergei Maslov, Ahmed Elbanna, Alexei Tkachenko, and Tong Wang—who graciously welcomed me into the modeling group and treated me as an equal. My thanks go to all of them and especially to my pre-COVID office mate, Zach Weiner, who graciously offered to join the COVID team and whose camaraderie and expertise over the past five years have kept me sane during many formative experiences.

My research career began at New York University in Andy Kent's spintronics lab, where I learned the importance of keeping detailed, concise reference notes. The next stop on my academic journey was in the field of wireless communications, where I had the opportunity to work with Ted Rappaport and Dennis Shasha on some very interesting projects that

ossified my irregular sleep schedule; it was Dennis's stimulating Freshman seminar that first encouraged me to pursue computational work. I found my way to astrophysics when Andrew MacFadyen gave me the opportunity to study relativistic hydrodynamic jets. I spent my time in Andrew's group under the attentive tutelage of Geoff Ryan, who donated hours of his time to help me debug my code, and whose enthusiasm is to blame for my original excitement over computational fluid dynamics. I am grateful to all of my professors and friends at NYU, who collectively helped establish me on my path to graduate school. I especially acknowledge Allen Mincer, with whom I had two short conversations that served as a disproportionately strong source of encouragement during the graduate school application process.

I thank Gil Holder, Lance Cooper, and Xin Liu for their thoughtful, guiding questions, and for their willingness to serve on my committee and devote some of their valuable time to my research world. My time in the physics department at Illinois has been greatly enhanced by Wendy Wimmer, whose manifest dedication to and support of its members has made my interactions with the university bureaucracy transparent. Finally, thank you again to Ben Prather, Vedant Dhruv, and Jonathan Ha for offering to review this document and for providing extensive feedback that undoubtedly improved its content and readability.

My graduate career at Illinois has been funded by NSF grants AST 17-16327 and OISE 17-43747, departmental support, and a Donald C. and F. Shirley Jones Fellowship. Funding for my time in New Mexico was provided by Los Alamos National Laboratory, which is operated by Triad National Security, LLC, for the National Nuclear Security Administration of the US Department of Energy (Contract No. 89233218CNA000001). My work has relied on various instruments, including the Chandra X-ray Observatory, the Nuclear Spectroscopic Telescope Array, and the many observatories that comprise the Event Horizon Telescope. I acknowledge the significance that Maunakea, the location of two EHT sites, has for the indigenous Hawai‘ian people. Many of the HPC resources I used were provided by NSF-supported XSEDE resources at the Texas Advanced Computing Center at The University of Texas at Austin.

To my mother.

Table of Contents

List of Tables

List of Figures

»Wahrlich es ist nicht das Wissen, sondern das Lernen, nicht das Besitzen, sondern das Erwerben, nicht das Da-seyn, sondern das Hinkommen, was den grössten Genuss gewährt.«

—Carl Friedrich Gauß

Chapter 1

Introduction

After observing Io for several years, Giovanni Domenico Cassini and Ole Rømer used the varying time delay between subsequent eclipses of the Jovian moon to infer the finite speed of light in 1676. Years later, near the end of the eighteenth century, John Michell and Pierre Laplace independently postulated the existence of *dark stars*, whose gravitational escape velocities would exceed the proposed speed of light, rendering them invisible. Studying optics and color, Huygens (1690) proposed the first mathematical wave model for light, but it received little support compared to the corpuscular theory of light often attributed to René Descartes and championed by Isaac Newton. Success of the Huygens wave model hinged on the existence of a so-called *luminiferous æther* that would serve as a medium through which the wave could propagate. Forward nearly two centuries, James Clerk Maxwell invoked the æther to model Michael Faraday's magnetic lines of force and ultimately developed his theory of electromagnetism (Maxwell, 1865).

The acceptance of dark stars diminished as the particle theory of light was supplanted by the theory of massless electromagnetic waves. But in the decades leading up to the twentieth century, a series of increasingly detailed experiments seemed to refute the existence of the theory's requisite æther. Historically, the æther had been invoked as an absolute reference frame, and as it fell out of favor, an apparent paradox in Maxwell's theory grew more dire: an observer will measure the speed of receding light as constant, irrespective of their own motion. So what would you see if you rode on a light beam?

Albert Einstein addressed this question with his special theory of relativity (Einstein, 1905), where he presented a framework of Lorentz invariance in which speed-through-space

could be traded for speed-through-time to allow both the observer-measured speed of light to be constant and the laws of physics to hold regardless of the observer's motion. Over the following decade, Einstein extended his framework to treat gravitation, culminating in his general theory of relativity (Einstein, 1916), in which gravity was no longer an instantaneous attractive force but rather the natural byproduct of curvature in a combined *spacetime*.

General relativity (GR) is written in the language of differential geometry. The curvature properties of the spacetime manifold are described by the *metric tensor*—a mathematical object that takes two vectors as input and outputs a scalar from which one can synthesize angles and distances. The metric tensor describes spacetime; GR defines how to construct the metric tensor via the Einstein field equations, which connect distributions of matter and energy to curvature. It is with this knowledge that we begin our journey.

1.1 Black holes

It is conventional to represent a spacetime metric by its *line element* $\mathrm{d}s$, which describes an infinitesimal displacement

$$\mathrm{d}s^2 \equiv g(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{x}) = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu. \tag{1.1}$$

Here, the Greek indices in the right-most term have been used to write the expression in general coordinates x^μ, $g_{\mu\nu}$ are the components of the covariant metric, and $\mathrm{d}x^\mu$ are the differentials for the coordinates. I use the mostly pluses convention for the metric signature $(-+++)$ and write in units where Newton's gravitational constant and the speed of light are unity, $G = c = 1$, unless otherwise specified.

The first exact solution to the Einstein field equations was derived by a professor at Göttingen, Schwarzschild (1916), during his time as German a soldier on the Russian front in World War I. Schwarzschild's solution describes spacetime around a non-spinning point

mass M. The Schwarzschild line element is

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2M}{r}\right)^{-1}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right), \tag{1.2}$$

where we have used a coordinate chart in which t is time, r is radius, θ is colatitude (or polar angle), and ϕ is azimuthal angle.

At $r = 0$, the coefficient of $\mathrm{d}t^2$ in Equation 1.2 diverges and the curvature of spacetime becomes infinite. At $r = 2M \equiv r_{\mathrm{eh}}$, however, the apparent divergence of the line element is an artifact of our coordinate choice: Lemaître (1933) showed it could be removed by re-expressing the metric in a different coordinate system. Yet r_{eh} is still interesting, since the radial and temporal components of the metric change sign at this radius, and the region of spacetime within $r = r_{\mathrm{eh}}$ is thus causally disconnected from its complement. Practically, the r_{eh} surface acts as both a one-way membrane and an edge to the visible universe. This boundary of no return is known as an *event horizon* (Finkelstein, 1958). We shall call an object that has an event horizon a *black hole*.

Extensions to the Schwarzschild metric followed. The metric for a charged point mass was found by Reissner (1916). Several decades later, a solution for a point mass with angular momentum was found by Kerr (1963), and Newman et al. (1965) found a solution for a point mass with both charge and angular momentum. Although Israel (1967, 1968) showed that the two non-rotating solutions were the only zero-angular-momentum black hole solutions, the uniqueness of the spinning solutions was found to be subject to several physically reasonable assumptions (see Carter (1979) for a review).

Astrophysical black holes are unlikely to have gravitationally significant charge,[1] but they

[1]The comparative strength of the electromagnetic force compared to gravity suggests a rapid charge-neutralization timescale, with the oppositely charged particles being sourced by either the local environment or electron–positron pairs produced by the black hole's electromagnetic field (see, e.g., Blandford & Znajek 1977).

may spin. The Kerr line element, which describes the uncharged spinning case, is

$$\begin{aligned} \mathrm{d}s^2 = &-\left(1-\frac{2Mr}{\Sigma}\right)\mathrm{d}t^2 - \frac{4Mra\sin^2\theta}{\Sigma}\mathrm{d}t\,\mathrm{d}\phi + \frac{\Sigma}{\Delta}\mathrm{d}r^2 + \Sigma\mathrm{d}\theta^2 \\ &+\left(r^2+a^2+\frac{2Mra^2}{\Sigma}\sin^2\theta\right)\sin^2\theta\mathrm{d}\phi^2, \end{aligned} \tag{1.3}$$

where the black hole angular momentum is written in terms of its (not dimensionless) spin parameter $a \equiv J/M$, $\Sigma \equiv r^2 + a^2\cos^2\theta$, and $\Delta \equiv r^2 - 2Mr + a^2$, and where we have used the coordinates described by Boyer & Lindquist (1967, hereafter BL), which are a generalization of Schwarzschild coordinates.

The Kerr metric has two radii at which the $\mathrm{d}r^2$ coefficient in the line element is singular

$$r_{\mathrm{eh},\pm} = M\left(1 \pm \sqrt{1-(a/M)^2}\right). \tag{1.4}$$

The outer horizon $r_{\mathrm{eh},+}$ is the analogue of the Schwarzschild event horizon in that it acts as a one-way membrane. Mathematically, the inner horizon is an (unstable) Cauchy horizon, within which geodesics may return to the same event and the foundations of causality fail. The physical spacetime singularity of the Kerr solution lies in the $\theta = \pi/2$ midplane at $r = 0$. Unlike in Schwarzschild, we find by integrating area over the $r = 0$ surface that $r = 0$ corresponds to the finite areal radius $r_{\mathrm{areal}} = \sqrt{r^2+a^2} = a$. By transforming to so-called Kerr–Schild Cartesian coordinates, with

$$x = (r\cos\phi + a\sin\phi)\sin\theta, \tag{1.5}$$

$$y = (r\sin\phi + a\cos\phi)\sin\theta, \tag{1.6}$$

$$z = r\cos\theta, \tag{1.7}$$

we see that the singularity takes the form of a ring at $z = 0$ with $x^2 + y^2 = r^2 + a^2 = a^2$.

We can find the "rotation" of spacetime near the black hole by taking the ratio of two metric components: in the Kerr metric, an inertial reference frame rotates with angular

velocity

$$\Omega = -\frac{g_{t\phi}}{g_{\phi\phi}}. \tag{1.8}$$

Evidently, the black hole drags spacetime with it as it spins. This rotational *frame dragging*, also known as the Lense & Thirring (1918) effect, has an interesting consequence in that it allows spin energy to be extracted from the black hole.

Frame dragging causes objects near a spinning black hole to appear to rotate from the perspective of a far-away observer—equivalently, if an object near a spinning black hole is to appear stationary to an observer at infinity, it must orbit the hole. The frame dragging effect is amplified at smaller radii and higher spins. The boundary beyond which objects must move faster than the speed of light in order to appear stationary is known as the ergosurface, within which lies the *ergoregion* (Ruffini & Wheeler, 1971). Although objects that enter the ergoregion are forced to corotate with the hole, they may ultimately escape its influence, since the ergoregion extends beyond the event horizon. Penrose & Floyd (1971) showed that a careful arrangement of matter could lead the dynamical spacetime to siphon angular momentum away from the hole and energize the escaping matter.

1.2 Black holes in nature

The first observational evidence for black holes came with the discovery of extragalactic quasars. Appearing at first to be merely faint, quasi-stellar objects, quasars were found in fact to be incredibly powerful. By identifying the optical counterpart of the radio source known as 3C273, Schmidt (1963) measured its redshift and used its distance to compute a bolometric luminosity, finding that the object was orders of magnitude more luminous than some galaxies. Subsequent measurements confirmed the result and found evidence of other, similar objects (e.g., Greenstein 1963; Matthews & Sandage 1963; Oke 1963; Hazard et al. 1963; Greenstein 1963; Greenstein & Matthews 1963; Greenstein & Schmidt 1964; Webster

& Murdin 1972). It was soon proposed that quasars could be powered by the accretion of matter onto supermassive black holes that were in excess of millions or billions of times the mass of our sun $M_\odot$ (Salpeter, 1964; Hoyle, 1966; Woltjer, 1966; Hoyle & Burbidge, 1966; Lynden-Bell, 1969; Soltan, 1982; Rees, 1984). Supermassive black holes are now thought to reside at the centers of almost all galaxies (see, e.g., Kormendy & Richstone 1995).

Observations of X-ray sources provide evidence for a class of smaller, stellar-mass black holes. The first case was provided by Cygnus X-1, an X-ray source in the constellation Cygnus that was discovered during a sub-orbital rocket flight (Bowyer et al., 1965). Optical observations of Cygnus X-1 revealed a variable blue supergiant with a hidden companion in a binary system. By measuring the velocity of the supergiant, the mass of its companion was estimated to be too large for a neutron star (see originally Bolton 1972; Webster & Murdin 1972 and the most recent mass estimate $\approx 21\ M_\odot$ from Miller-Jones et al. 2021). The black hole companion theory provided an opportune explanation of the X-ray emission as the product of mass accretion from the supergiant onto the hole. Binary star systems are common, and every galaxy may contain tens to hundreds of millions of stellar-mass black holes (see Remillard & McClintock 2006 for a review).

The Milky Way is likely filled with stellar-mass black holes, but the nearest supermassive black hole candidate lies at its center. The object was first noticed when Balick & Brown (1974) identified a strong, compact radio source near the constellation Sagittarius.[2] The radio source became known as Sagittarius A following the naming convention, and Brown (1982) labeled the exciting compact feature Sgr A* (read Sagittarius A-star). It is difficult to observe Sgr A* in the optical- and ultraviolet-wavelengths because of the large column of gas and dust along the line of sight through the galactic plane. Nevertheless, its relative closeness has encouraged numerous observational campaigns across the electromagnetic spectrum from the radio to γ ray (e.g., Goldwurm et al. 1994; Merck et al. 1996; Falcke

[2]Lynden-Bell & Rees (1971) had predicted the existence of a supermassive black hole at the center of our galaxy several years earlier.

et al. 1998; Genzel & Eckart 1999; Cotera et al. 1999; Genzel et al. 2003; Baganoff et al. 2003; An et al. 2005; Schödel et al. 2007; Bower et al. 2015; Liu et al. 2016b,a; Schödel et al. 2011; Witzel et al. 2012; Neilsen et al. 2013). The radio spectrum is almost flat and peaks at millimeter wavelengths. Flares in the X-ray tend to occur simultaneously with near-infrared (NIR) flares, and the variability timescale of certain flaring events constrains the size of the variable emitting region to be of order the radius of the black hole event horizon (see, e.g., Dodds-Eden et al. 2009)

General interest in Sgr A* and its closeness to Earth have inspired detailed study of its neighborhood. Early measurements of stellar and gas emission lines suggested the presence of a $\sim$ million-solar-mass object within the compact radio source (Wollman et al., 1977; Lacy et al., 1980, 1982; Haller, 1992; Genzel et al., 1996). Technological advances, including adaptive optics (see Beckers 1993), have enabled precise tracking of stellar orbits and gas motion at micro-arcsecond scales. Models of this motion agree with the emission line measurements and are consistent with the presence of a localized mass $\approx 4 \times 10^6\ M_\odot$ at a distance ≈ 8.3 kpc (see, e.g., Schödel et al. 2002; Ghez et al. 2004, 2008; Gillessen et al. 2012, 2017; Gravity Collaboration 2017). Its small size and rapid variability strongly suggest that Sgr A* is a supermassive black hole. Recent high-precision astrometry measurements of stellar orbits near Sgr A* show evidence of a perihelion precession consistent with the prediction of a black hole in GR (Gravity Collaboration et al., 2020a).

By balancing an object's inward gravitational pull on nearby matter with an outward radiation pressure, the characteristic luminosity associated with an object of mass M can be determined. If the surrounding matter comprises ionized hydrogen, then the so-called Eddington luminosity is

$$L_{\rm Edd} = \frac{4\pi G M m_p c}{\sigma_T}, \tag{1.9}$$

where m_p is the mass of a proton and σ_T is the Thomson cross section (see, e.g., Rybicki

& Lightman 1979). Compared to its Eddington luminosity, Sgr A* is incredibly dim ($\lesssim 10^{-6}\ L_{\rm Edd}$, see Narayan et al. 1998 and references within). Sgr A* thus belongs to a class of low-luminosity active galactic nuclei (LLAGN). Sgr A*'s large mass and close proximity establish its unique position as the largest known object of its class as measured by angular extent on the sky.

The nearby giant elliptical galaxy Messier 87 (also known as Virgo A or NGC 4486) hosts an LLAGN with a radio flux comparable to that of the Sgr A* source, but its claim to fame was its relativistic jet. Curtis (1918) described the galaxy[3] as "[e]xceedingly bright" with "[a] curious straight ray ... apparently connected with the nucleus by a thin line of matter." After measuring the polarization properties of the ray, Baade & Minkowski (1954) theorized that the emission was produced by electrons moving at relativistic speeds and were the first to identify the feature as an astrophysical relativistic jet. Observations at increasing frequencies have found that the jet structure persists into the bright radio core of the galaxy at sub-parsec scales (e.g., Palmer et al. 1967; Reid et al. 1982; Junor et al. 1999; Kovalev et al. 2007; Ly et al. 2007; Asada & Nakamura 2012). Kinematic modeling and long-term monitoring have established that the inclination angle between the jet axis and the line of sight is $\approx 20^\circ$ (e.g., Heinz & Begelman 1997; Mertens et al. 2016; Walker et al. 2018) and that the kinetic energy within the jet is within the range $10^{42} - 10^{45}$ erg s^{-1} (e.g., Owen et al. 2000; Stawarz et al. 2006; de Gasperin et al. 2012).

Studies of the M87 host galaxy infer the existence of a central supermassive black hole. By fitting dynamical models of star clusters to observed radial velocity profiles and velocity dispersions, Sargent et al. (1978) deduced that $\approx 5 \times 10^9\ M_\odot$ of mass must be concentrated in the center of M87 and noted the possibility that the mass was due to a supermassive black hole. Recent stellar dynamics measurements refined the mass estimate to $\approx 6.2 \times 10^9\ M_\odot$

[3]The existence of distinct galaxies beyond our own was conjectured hundreds of years ago, but the idea did not gain widespread acceptance until the 20th century. Accordingly, Messier (1781) classified M87 as a "Nebula without star."

(Gebhardt et al., 2011), although they are in disagreement with the $\approx 3.5 \times 10^9\ M_\odot$ estimate from models of gas dynamics near the core (Walsh et al., 2013). High resolution imaging and modeling by the Event Horizon Telescope (EHT) collaboration produce mass estimates that agree with the stellar dynamics figure at $\approx 6.5 \times 10^9\ M_\odot$ (Event Horizon Telescope Collaboration et al., 2019a). The M87 and Sgr A* black hole candidates are the two largest known black holes on the sky, and in a cosmic coincidence, the ratio between their masses is nearly the ratio between their distances from Earth.

1.3 Accretion and the black hole–jet connection

The M87 jet is not unique. In fact, it appears that astrophysical relativistic jets are common, and they seem to be launched by active galactic nuclei (see Blandford et al. 2019 for a review). Where does their energy come from? Evidently the engines must be capable of sustained power output over at least thousands or millions of years. In M87, the apparent co-location of the jet base and the putative central supermassive black hole suggests a possible connection between the two. We begin with a review of accretion theory to explore the possibility that the black hole drives the jet.

Although black holes capture everything that passes through their event horizons, black hole accretion systems are among nature's most efficient mechanisms for turning mass into energy. The conversion is an implicit consequence of disk accretion. As matter falls toward a massive object from infinity, its gravitational potential energy transforms to kinetic energy; if the matter is to follow a nearly circular path to the black hole, it must liberate some of this kinetic energy in order to maintain the correct velocity at each radius. In the astrophysical context, black hole accretion flows are composed of ionized plasma from stellar winds, gas clouds, &c. In these gaseous flows, local interactions may convert the kinetic energy into thermal energy (e.g., via viscous dissipation, turbulent cascades, and shocks). The gas heats, and the accretion disk glows.

In addition to liberating kinetic energy, angular momentum must be transported outward in order for accretion to proceed, since the angular momentum of matter in a circular orbit increases with radius. Shakura & Sunyaev (1973), and Novikov & Thorne (1973) in the relativistic case, proposed that viscosity could be responsible for angular momentum transport and derived the first analytic disk models from this prescription (see also Pringle & Rees 1972). The disks in these models have high radiative efficiencies and cool to a geometrically thin state. If the disks cannot cool quickly enough, they are said to be advection-dominated accretion flows (ADAF, also see radiatively inefficient accretion flow RIAF, Ichimaru 1977; Rees et al. 1982; Narayan & Yi 1994, 1995a; Abramowicz et al. 1995). In an ADAF/RIAF, the disk remains hot as the energy is carried with the flow and ultimately lost through the event horizon. ADAFs are geometrically thick and have luminosities that are small compared to $L_{\rm Edd}$.

The details of the angular momentum transport were not well understood until the rediscovery of the magneto-rotational instability (MRI, Balbus & Hawley 1991), in which magnetic field lines connecting plasma at different radii in a differentially rotating flow torque smaller (larger) radii to slower (faster) speeds. The onset of the MRI was shown to foster the development of turbulence in shearing box simulations (e.g., Hawley et al. 1995; Balbus & Hawley 1998). Evidently, magnetic fields may play an important role in enabling disk accretion.

In the astrophysical black hole accretion scenario, the conductivity of the accreting plasma is high. Alfvén (1942) showed that in the limit of infinite conductivity, magnetic field lines are "frozen in" the plasma—any plasma motion perpendicular to the direction of the field lines would generate infinite eddy currents and is thus disallowed. What, then, happens as magnetic field lines are advected toward the event horizon?

If the accretion flow revolves about the hole as it falls, then the magnetic field lines will wind around the spin axis of the system. When the black hole is spinning, Blandford & Znajek (1977, BZ) showed that the ergoregion would drive an extra torque on the field lines in a

process that extracts the spin energy of the hole and converts it into an outward electromagnetic Poynting flux jet aligned along the angular momentum axis. BZ jet power scales like the square of both the black hole spin and the magnetic flux on the horizon. Numerical fluid simulations of black hole accretion flows have been shown to produce electromagnetic jets that are consistent with the BZ model (see, e.g., McKinney & Gammie 2004; Tchekhovskoy et al. 2011).

1.4 Direct observations of black holes

Although there is myriad evidence supporting the existence of black holes, recent technological advances have enabled significant strides in observational techniques. The past six years have seen a rapid increase in measurement precision and fidelity, allowing us to investigate black holes directly.

One method for direct detection relies not on light but gravity. When massive objects accelerate through space, GR predicts they will produce a disturbance in the underlying spacetime in the form of gravitational waves (Einstein 1916 and see also Poincaré 1906). When two compact objects merge, the spacetime disturbance is great, and an abundance of gravitational waves will be produced. Although deviations from the precise periodicity of pulsars signals has been cited as evidence of gravitational waves for decades (Hulse & Taylor 1975 and see, e.g., Hobbs et al. 2010), the first direct detection of a gravitational wave was made by the Laser Interferometer Gravitational-Wave Observatory (LIGO) in 2015 (Abbott et al., 2016a). Since the initial detection, LIGO has observed multiple black hole mergers and neutron star mergers, enabling precise tests of GR (Abbott et al., 2016b, 2017, 2019a). LIGO results indicate that stellar mass black hole mergers are frequent, at a rate of $10 - 100$ Gpc^{-3} yr^{-1} (Abbott et al., 2019b).

The angular resolution of a telescope is proportional to the ratio of its observing wavelength λ to its size d, so the relatively small angular size of black holes makes it a challenge

to image them directly. Using a mass of $4 \times 10^6\ M_\odot$ at a distance of 8000 pc for the putative Sgr A* black hole, we find that one gravitational radius GM/c^2 subtends $\approx 5\ \mu$as on the sky. Gravitational lensing deflects the geodesic paths of light near a black hole, causing some light rays to fall onto the event horizon. The boundary between the geodesics that are captured by the hole and those that escape sets a characteristic black hole "shadow" radius on the image. The shadow radius $= \sqrt{27}\,GM/c^2$ if the black hole has no angular momentum; it is slightly smaller otherwise (see, e.g., Hilbert 1917; Darwin 1959; Bardeen 1973; Luminet 1979; Ohanian 1987; Takahashi 2004). The diameter of Sgr A*'s shadow is $\approx 50\ \mu$as. At $\lambda = 1.3$ mm, the angular resolution of a telescope the size of the earth is $\approx 25\ \mu$as, putting shadow-scale imaging of both Sgr A* and M87 just within reach.[4]

The Event Horizon Telescope is a heterogeneous global network of millimeter and submillimeter observatories that uses very long baseline interferometry (VLBI) techniques to reconstruct high resolution images of radio sources (Doeleman et al., 2009). VLBI techniques have been used for decades (see Kellermann & Moran 2001 for a review) and operate on the principle of correlating signals and timing data recorded at different observatories in order to perform interferometry.

An interferometer produces data in the form of "complex visibilities," i.e., the correlation between the electromagnetic field measured at a pair of observatories. Each visibility is defined at a baseline, which is determined by the projected distance and orientation of the station-to-station separation as measured on the $u - v$ plane perpendicular to the line of sight. The visibilities are the Fourier transform of the image intensity distribution, so longer baselines provide higher resolutions. Because radio telescopes do not cover the Earth, the Event Horizon Telescope only populates the limited portion of the $u - v$ plane swept out by pairs of participating telescopes as the Earth rotates over the course of the observation.

The process of translating the sparse Fourier data into an image is known as image re-

[4]As noted above, the absolute size of a black hole scales linearly with its mass M, and its angular size thus goes like the ratio between its mass and distance D. Since M87 and Sgr A* have approximately equal M/D, they are approximately the same size on the sky.

construction. The EHT collaboration used both the conventional CLEAN method (Högbom 1974; Clark 1980; the EHT uses `DIFMAP`, Shepherd 1997) and the newer regularized maximum likelihood technique (`eht-imaging`, Chael et al. 2016, 2018b, and `SMILI`, Akiyama et al. 2017a,b) to perform image reconstruction. Validation of the image reconstruction procedure is described in Event Horizon Telescope Collaboration et al. (2019d). The first EHT results in total intensity (Event Horizon Telescope Collaboration et al., 2019a,b,c,d,e,f) and linear polarization (Event Horizon Telescope Collaboration et al., 2021a,b) showed consistency between the source image and the predicted image of a black hole in GR. They also produced model constraints on black hole mass, accretion rate, spin, and the local strength of the magnetic field near the event horizon.

The GRAVITY instrument is another interferometer comprising the four telescopes of the Very Large Telescope operated by the European Space Agency in Chile. GRAVITY operates in the NIR band between $\lambda \approx 2 - 2.4$ μm, giving it a spatial imaging resolution of ≈ 4 mas. GRAVITY also has the ability to perform astrometry measurements of positions and motion with a precision on the order of tens of μas (Gravity Collaboration, 2017). The GRAVITY collaboration identified circular motion at $30 - 50\,\mu$as scales during NIR observations of Sgr A* in its flaring state over the course of an hour. The measurements were consistent with the idea that the flare could be due to a hot spot orbiting the hole at a very high viewing angle between $140 - 160°$, (Gravity Collaboration et al., 2018). GRAVITY measurements of linear polarization were also used to infer the structure and strength of magnetic fields near the event horizon (Gravity Collaboration et al., 2020b).

1.5 Dissertation outline

The projects described in this dissertation are motivated by the goal of probing the connection between black holes and relativistic jets via observation. Recent and continuing advances in technology have enabled high-precision horizon-scale imaging of black hole accretion flows.

Current and next generation observations may be able to validate the theory that relativistic jets like the one in M87 are powered by supermassive black holes via the Blandford–Znajek mechanism, but obtaining unambiguous evidence requires a thorough understanding of the observational signatures of black hole spin and horizon-scale magnetic field structure. My dissertation covers both my contributions to the observational analysis of the M87 black hole performed by the Event Horizon Telescope collaboration as well as my predictive modeling of astrophysical processes near the hole, including studies of the horizon-scale jet composition and a novel universal signature of black hole spin.

I begin in Chapter 2 with a review of accretion systems and radiative physics in the astrophysical context. I then describe the numerical codes I use to generate fluid simulations and synthesize electromagnetic observables. This description includes a review of the `iharm`, `ipole`, and `igrmonty` codes and focuses on the science-enabling modifications I have made to the code bases. I conclude with a brief example of simulation data products and a discussion of in-progress and near-future ways to improve the simulation pipeline.

Chapter 3 describes my contributions to the global Event Horizon Telescope collaboration effort to provide a theoretical interpretation of the initial 2019 total intensity image of the M87 black hole, which produced bounds on the mass and spin orientation of the putative supermassive black hole. Chapter 4 continues discussion of my work with members of the EHT collaboration and begins with a description of the development, testing, and validation of a new polarimetric observable—the PWP β metric—that can be used to constrain the magnetic field structure near the black hole event horizon. It concludes with an overview of the 2021 EHT analysis of linear polarized image data from the M87 source.

The last part of my dissertation is more theory-based and begins in Chapter 5 with a discussion of matter entrainment at the jet–disk boundary region between the accretion flow and the outflowing relativistic jet, which reaches down to horizon scales near the black hole. In this chapter, I show that the boundary layer is episodically unstable to the magnetic Kelvin–Helmholtz instability. Using numerical simulations, I find evidence of the instability's

operation and show that the mass entrainment rate from the disk to the jet may reach $\approx 1\%$ the accretion rate through the event horizon.

In Chapter 6, I continue consideration of mass within the jet, this time focusing on the electron–positron pairs that may be produced by photon–photon interactions from the $\sim 1\,\mathrm{MeV}$ component of the background radiation field. I find that the charge density sourced by this drizzle pair production process may be high enough to short out unscreened electric fields and prevent the pair cascade process from operating in the jet. The composition of the radiating leptons in the jet has observational consequences, especially in the context of circular polarization measurements.

Finally, in Chapter 7, I describe my novel universal black hole angular momentum observable, black hole glimmer, which comprises a position-dependent autocorrelation signature in black hole movies. I provide a detailed analytic description of black hole glimmer, which is due to the geometric structure of nearly bound photon orbits near spinning black holes, and I finish with several example measurements and a demonstration of how the glimmer signature is independent of the properties of the emitting plasma.

Chapter 2

Simulating electromagnetic black hole observables

In this chapter, I discuss the procedure I use to generate synthetic electromagnetic observables of supermassive black hole accretion systems. After a brief motivating introduction, I present a theoretical background describing the target accretion systems, radiative physics, and the equations that our numerical tools solve. I then survey the details of the three codes I use most often in my work, paying particular attention to modifications and science-enabling enhancements that I have made to them. After the code description, I outline my contribution to the simulation library used by the Event Horizon Telescope collaboration analysis of M87 and provide several example figures that demonstrate details of the simulation pipeline. I end with a discussion of limitations and future directions for the simulation pipeline.

2.1 Introduction

The Event Horizon Telescope (EHT) is a globe-spanning network of millimeter wavelength observatories that is capable of imaging nearby supermassive black holes at event-horizon scale resolutions. In 2019, the EHT published the first total intensity images of horizon-scale emission from the center of the giant elliptical galaxy Messier 87 (Event Horizon Telescope Collaboration et al., 2019a,b,c,d,e,f, hereafter EHTC I–VI), and in 2021, they published linear polarization data on the same source (Event Horizon Telescope Collaboration et al., 2021a,b, hereafter EHTC VII–VIII). Data from observations of the galactic center, Sgr A*, are expected soon. Improvements to the EHT including extra interferometric baselines,

Parts of this chapter are in preparation for submission to The Astrophysical Journal Supplement Series.

increased sensitivities, and support for measurements at new frequencies indicate that more detailed measurements lie on the horizon.

The EHT analysis relies in part on a library of synthetic observations. This library comprises a large set of numerical simulations of magnetized, relativistic black hole accretion flow models and accompanying polarized ray-traced simulated images and spectral energy distributions. Synthetic observations serve two primary functions: they can be used to validate image reconstruction procedures, and they can be used in the forward modeling pipelines that compare and contrast the observational and synthetic data to infer the physical parameters of the system.

General relativistic magnetohydrodynamics (GRMHD) simulations are a mainstay in analysis of black hole accretion flows (e.g., Gammie et al. 2003; Mignone et al. 2007; Del Zanna et al. 2007; Narayan et al. 2012; Sądowski et al. 2013b; White et al. 2016; Porth et al. 2017; Liska et al. 2018), and many general relativistic ray-tracing codes have been developed for both imaging (Noble et al., 2007; Psaltis & Johannsen, 2012; Chan et al., 2013, 2018; Mościbrodzka & Gammie, 2018; Dexter et al., 2020) and the generation of spectra (Dolence et al., 2009; Mościbrodzka et al., 2009; Dolence et al., 2012; Zhang et al., 2019; Mościbrodzka, 2020).

In this chapter, I describe the modeling pipeline `PATOKA`, which was used for the Illinois contribution of fluid accretion models, images, and spectra to the EHT library. I here describe differences between the published code descriptions and the versions of code that were used in generating the library. In Section 2.2, I provide a brief summary of the theoretical underpinnings of the accretion model and radiative physics. In Section 2.3, I review the details of the numerical codes used to perform the simulations and ray tracing. In Section 2.4, I present some results that were generated with the pipeline. Numerical details are described in Sections 2.6 and 2.7. I end with a discussion of future directions in Section 2.5.

2.2 Theoretical background

We begin with a brief overview of the theoretical considerations associated with library generation. We cover the parameters of the black hole accretion model and review the equations governing radiative transfer in hot astrophysical plasmas.

2.2.1 Black hole accretion

The supermassive black holes at the center of M87 and the galactic center (Sgr A*) are associated with compact radio sources (Doeleman et al., 2008; Fish et al., 2011) and serve as the primary targets for the EHT. M87 and Sgr A* are low luminosity active galactic nuclei (see Greene & Ho 2007; Yuan & Narayan 2014 for reviews), so the accretion onto their central black holes is expected to proceed at a low rate $\dot{m} \equiv \dot{M}/\dot{M}_{\rm Edd} < 10^{-3}$ and therefore be radiatively inefficient (RIAF; Ichimaru 1977; Narayan & Yi 1994, 1995b; Quataert & Narayan 1999; Yuan et al. 2003). In RIAF models, the gravitational energy of the flow, which must be shed during accretion, is either advected across the event horizon or lost in mechanical outflows. RIAFs comprise geometrically thick disks of relativistically hot, magnetized plasma (Rees et al., 1982; Narayan & Yi, 1995b; Narayan et al., 1995; Yuan & Narayan, 2014; Reynolds et al., 1996; Yuan et al., 2002; Di Matteo et al., 2003). Advection-dominated accretion flows are typically collisionless, and since the electrons and ions are subject to different heating mechanisms, the flows are likely two-temperature (see Shapiro et al. 1976; Mahadevan & Quataert 1997; Quataert 1998; Ressler et al. 2015). M87 also supports a FR-I type relativistic jet that extends to kiloparsec scales and has an estimated jet power of $\approx 10^{42} - 10^{44}$ erg/s (Stawarz et al., 2006; de Gasperin et al., 2012; Prieto et al., 2016).

Although analytic models are useful in understanding the broad dynamics of accretion physics (see especially the early models of Novikov & Thorne 1973; Shakura & Sunyaev 1973), the nonlinearity of the magnetohydrodynamics equations fosters the development of

shocks and turbulence that make a pure analytic treatment infeasible.[1] Since a large part of the millimeter emission produced in a RIAF flow is likely produced near the horizon (e.g., Mościbrodzka et al. 2009), a full general relativistic treatment of the system may be necessary to adequately recover the detailed features of the model.

Numerical, general relativistic magnetohydrodynamics (GRMHD) simulations serve as a more practical method for studying the plasma dynamics in detail. GRMHD simulations evolve the distribution of plasma and electromagnetic energy through time and produce a time series description of the plasma state. Global GRMHD simulations have been widely used to study RIAF accretion and have been shown to reproduce the jet powers of the Blandford & Znajek (1977) model (e.g., Koide et al. 1999; De Villiers & Hawley 2003; Gammie et al. 2003; McKinney & Gammie 2004; Narayan et al. 2012).

At low accretion rates, radiative cooling is negligible (see, e.g., Dibi et al. 2012; Ryan et al. 2017) and the evolution of the fluid is invariant under rescalings of both the metric length GM/c^2 and the mass density. Ideal GRMHD simulations thus only introduce the following physical parameters: the angular momentum of the system and the magnetic flux near the event horizon.

Angular momenta and tilted disks

The angular momentum of the central black hole J is often expressed in terms of the dimensionless black hole spin parameter $a_* \equiv Jc/GM^2$. The orientation of the accretion flow near the black hole is not necessarily set by the boundary processes that supply the system with plasma, and the angular momentum vectors of the black hole and the accretion flow may be misaligned with the black hole spin. The angular separation between the two axes is called tilt. Although there are plausible scenarios that produce accretion flows with zero tilt, there is at present no way of rejecting models with strong or even maximal (180 degree,

[1]The magnetorotational instability of Balbus & Hawley (1991) may play a crucial role in facilitating the angular momentum transport necessary for accretion in steady disks, whose horizon-scale flows are not choked, cf. MAD flows in §2.2.1.

i.e., retrograde) tilt.

Systems with non-zero tilt may have clear observational signatures. Fragile & Blaes (2008); Dexter & Fragile (2013) found that tilted accretion flows could produce a two-arm shock that would significantly alter the morphology of radio images and the shape of the spectrum. Some simulations suggest tilted disks ultimately align with the central black hole via an analogue of the Bardeen & Petterson (1975) effect (McKinney et al., 2013; Liska et al., 2018).

In our galaxy, there is strong evidence that the central supermassive Sgr A* was recently in an active period of higher accretion (Totani, 2006; Ponti et al., 2013), and other observational studies and simulations report evidence of clumpiness within a few parsecs of Sgr A* in the circumnuclear disk (Montero-Castaño et al., 2009; Blank et al., 2016). Ressler et al. (2018, 2020a,b) simulated mass feeding in the Galactic Center via magnetized stellar winds and found that a wide variety of different initial configurations led to the development of similar non-disk-like accretion flows with strongly poloidal horizon-scale magnetic field configurations. Evidently, the angular momentum of the boundary condition that supplies the inner accretion flow with mass is likely to be variable on galactic timescales (e.g. Cuadra et al., 2006).

More generally, a non-trivial time-dependent boundary feeding condition is not unlikely since there is a large discrepancy between the long timescale that governs changes to properties of black hole and the shorter ones on which the local environment of the black hole (i.e., the environment that governs the its feeding) changes. Thus, the bulk angular momentum of the accreting matter (far from the horizon) may exhibit a time-dependent tilt with respect to the spin angular momentum of the central black hole.

Time-dependent tilt misalignments have been invoked to explain the wobble observed in some relativistic jets (see Natarajan & Armitage, 1999). These arguments find extra support in the analysis of episodic jet variations and spectral analyses of the broad Fe line (e.g., Hjellming & Rupen, 1995; Rout et al., 2020). Theoretical studies that consider the

initial spin distribution of supermassive black holes (De Luca et al., 2019) and simulate black hole growth via accretion and mergers in hierarchical galaxy formation models also often favor random alignment between the black holes and their disks (see, e.g., Volonteri et al., 2005).

The prograde/retrograde dichotomy has also been used to explain differences between radio-loud and radio-quiet active galactic nuclei sources (Garofalo 2009; Garofalo et al. 2010, but see Tchekhovskoy & McKinney 2012). Some observational studies have claimed detection of retrograde accretion systems: Morningstar et al. (2014), Chen et al. (2016), and Mikhailov et al. (2018) presented criteria for classifying systems as retrograde and singled out several known systems based on emission models and low estimated Eddington ratios.

Magnetic flux

The accreting plasma carries magnetic field lines with it as it falls onto the central black hole. If the sign of the magnetic flux is constant over time, then the field lines will accumulate and increase the magnetic pressure near the hole, eventually saturating when the magnetic pressure is large enough to counterbalance the inward ram pressure of the accretion flow. The amount of magnetic flux threading the event horizon thus qualitatively divides accretion flows into two categories: the magnetically arrested disk (MAD) state (Bisnovatyi-Kogan & Ruzmaikin, 1974; Igumenshchev et al., 2003; Narayan et al., 2003), in which magnetic pressure is large and the magnetic field is dynamically important, and the alternate standard and normal evolution (SANE) state (Narayan et al., 2012; Sądowski et al., 2013a).

Quantitatively, the magnetic flux through the horizon is measured in terms of an integral over a hemisphere at the event horizon

$$\Phi_{\rm BH} = \frac{1}{2}\int_{r=r_{\rm eh}} \left|{}^{\star}F^{rt}\right| \sqrt{-g}\,{\rm d}\theta{\rm d}\phi, \tag{2.1}$$

but it is conventionally written as ϕ, a quantity normalized by the mass accretion rate of

the system $\dot{M}$

$$\phi \equiv \frac{\Phi_{\rm BH}}{\sqrt{\dot{M} r_g^2 c}}, \tag{2.2}$$

where $r_g \equiv GM/c^2$ is the gravitational radius of the hole. In the conventional rationalized Lorentz–Heaviside units used in GRMHD simulations, ϕ saturates at $\phi_c \approx 15$ (see Tchekhovskoy et al. 2011 and Porth et al. 2019).[2]

MAD accretion is choppy and tends to proceed in isolated, thin plasma streams that begin far from the hole, and MAD flows are often punctuated by violent magnetic bubble eruption events that release excess trapped magnetic flux. Although flux ejection events are fully understood, they may be contextualized in terms of a magnetic Rayleigh–Taylor interaction between the disk and the hole (Marshall et al., 2018).

Simulations suggest that the MAD vs. SANE dichotomy is observationally encoded in signatures of polarization, variability, and the details of the jet–disk connection. Recent data and analysis from the EHT suggest that M87 accretion is MAD (EHTC V; EHTC VIII).

Electron temperature and radiative effects

GRMHD simulations typically treat the accreting plasma as a single, thermal fluid and track only the internal energy (or temperature) of the bulk plasma. Conventional methods for determining the electron temperature typically involve setting T_e from the local fluid parameters, such as the internal energy of the fluid and the local plasma β ratio of gas pressure to magnetic pressure $\beta \equiv P_{\rm gas}/P_{\rm mag}$ (e.g., Goldston et al. 2005; Mościbrodzka et al. 2009; Shcherbakov et al. 2012; Mościbrodzka et al. 2014; Mościbrodzka & Falcke 2013; Mościbrodzka et al. 2016—see Anantua et al. 2020b for a comparison of several different models).

[2]In Lorentz–Heaviside units, the vacuum permittivity and permeability are $\epsilon_0 = \mu_0 = 1$ like in Gaussian units, but the strength of the electric and magnetic fields (and the reciprocal of the fundamental charge) differ from Gaussian units by a factor of $\sqrt{4\pi}$. In Gaussian units, $\phi_c = 50$.

GRMHD simulations that recover the turbulent dynamics of a flow often rely on implicit large eddy simulations (ILES; see Boris 1990; Boris et al. 1992), wherein it is assumed that numerical dissipation at the smallest numerical resolutions emulates the effects of physics at scales smaller than the grid-scale. Ressler et al. (2015); Sądowski et al. (2017) described methods to track this numerical dissipation in relativistic magnetohydrodynamics simulations. Given a model for sub-grid electron heating, the dissipated energy can be distributed into the plasma ions and electrons and therefore used as a heating model to independently track and evolve the electron temperature of the fluid. Several such kinetic plasma physics models have been proposed (Howes, 2010; Rowan et al., 2017; Werner et al., 2018; Kawazura et al., 2019; Zhdankin et al., 2019).

At low accretion rates $\dot{M}$, radiative losses are too slow to materially alter the internal energy of the flow (see, e.g., Sharma et al. 2007; Dibi et al. 2012), so pure GRMHD simulations sufficiently recover the dynamics. As $\dot{M}$ approaches 10^{-5} of the Eddington rate, however, it becomes important to include synchrotron emission and Compton upscattering and to include the contribution of the radiation stress–energy tensor through the use of radiation–GRMHD schemes, like `KORAL` (Sądowski et al. 2013b, implementing a gray M1-closure scheme for the radiation), `ebhlight` (Ryan et al. 2015, 2019, using a full Monte Carlo treatment of the radiation field), or `MOCMC` (Ryan & Dolence 2020, using an adaptively refined Monte Carlo treatment of the radiation field). The transitory regime in which radiation becomes important has also been studied by, e.g., Fragile & Meier (2009); Wu et al. (2016); Sądowski & Gaspari (2017); Sądowski et al. (2017); Ryan et al. (2017).

2.2.2 Radiative transfer

230 GHz simulated images of radiatively inefficient accretion flows near supermassive black holes are dominated by a distinct ring-like feature with large brightness temperature. In simulations, the location of the ring is broadly consistent with the critical curve boundary within which all geodesics both pass through the camera plane and terminate on the black

hole event horizon. Geodesics close to but outside of the critical curve are still lensed near the hole. The set of all geodesics that are lensed enough to complete n complete half-orbits around the hole defines the nth photon subring (Johnson et al., 2020); as $n \to \infty$, the subrings approach the critical curve. Each subring corresponds to a delayed, demagnified image of the universe.

In astrophysical scenarios, the images are dominated by emission produced near the black hole. Summing of the self-similar demagnified subring images produces a sharp, narrowing feature in intensity whose peak converges to the critical curve in the limit that the optical depth of the system is negligible. The feature seen in the composite image is known as the "photon ring" of the black hole (Bardeen, 1973; Luminet, 1979; Gralla & Lupsasca, 2020; Johnson et al., 2020), and the region it encircles is the so-called black hole shadow. The location of the critical curve is controlled by properties of the underlying spacetime geometry. Various strategies to constrain the properties of the spacetime by measuring the critical curve via the shape of the photon ring or due to autocorrelations and subring image delays have been proposed (e.g., Falcke et al. 2000; Takahashi 2004; Bambi & Freese 2009; Hioki & Maeda 2009; Amarilla et al. 2010; Amarilla & Eiroa 2013; Tsukamoto et al. 2014; Younsi et al. 2016; Mizuno et al. 2018; Johnson et al. 2020; Medeiros et al. 2020; Olivares et al. 2020; Wielgus et al. 2020; Wong 2021; Hadar et al. 2021; Chesler et al. 2021).

Particle trajectories in curved spacetimes are determined by the geodesic equations

$$\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} = p^\alpha \tag{2.3}$$

$$\frac{\mathrm{d}p^\alpha}{\mathrm{d}\lambda} = -\Gamma^\lambda{}_{\alpha\beta} p^\alpha p^\beta, \tag{2.4}$$

where Γ is a Christoffel symbol, λ is an affine parameter, and where for a photon the momentum $p^\alpha \to k^\alpha$, the photon wavevector. Notice that the frequency ν of the photon as measured in a frame with four velocity u^α is given by $\nu = -k^\alpha u_\alpha$. The intensity of the light changes along the geodesics according to its interaction with the local matter in four ways:

emission can increase the intensity of the light, scattering into the geodesic can increase the intensity of light, the intensity can decrease either due to absorption or scattering out of the geodesic (collectively called extinction), and the local plasma properties can mix polarization modes

$$\mathrm{d}I = \mathrm{d}I_{\mathrm{em}} + \mathrm{d}I_{\mathrm{sc}} + \mathrm{d}I_{\mathrm{ext}} + \mathrm{d}I_{\mathrm{mix}}. \tag{2.5}$$

Here, the extinction term depends on the local intensity of the incoming light

$$\mathrm{d}I_{\mathrm{ext}} = -\beta_{\mathrm{ext}} I \,\mathrm{d}s \tag{2.6}$$

and an extinction coefficient β_{ext}, which depends on the local plasma parameters and is a function of the angle the light ray makes with the magnetic field in the plasma.

Polarization

In the supermassive black hole accretion systems targeted by the EHT, the observed 230 GHz radiation is due primarily to synchrotron radiation, which is produced when electrons in the accreting plasma are accelerated by magnetic fields (see, e.g., Yuan & Narayan 2014). The local orientation of the magnetic field sets a preferred direction for the electromagnetic radiation, and thus synchrotron emission is, in general, polarized according to the structure of the magnetic field.

The full description of the polarized state of the light is often written in terms of the specific intensities of the Stokes parameters I_ν, Q_ν, U_ν, V_ν. Here, I_ν is the familiar total specific intensity of the light; Q_ν and U_ν represent the linearly polarized intensities; and V_ν represents the circularly polarized one. Not all of the light must be polarized, so $I_\nu^2 \geq$

$Q_\nu^2 + U_\nu^2 + V_\nu^2$. In this representation,[3] the electric vector position angle (EVPA) χ is

$$\chi = \frac{1}{2} \arctan \frac{U}{Q}. \tag{2.7}$$

Physically, χ describes the angle the electric field oscillations make with respect to some fiducial coordinate system, and the factor of one-half is due to the π-fold symmetry in describing this orientation.

As polarized light travels through a magnetized plasma, two effects cause χ to rotate. The first is due to parallel transport of the polarization vector in the curved geometry. The magnitude of this effect is a pure function of the underlying spacetime (see Gelles et al. 2021a for a discussion). The second rotation is due to the so-called Faraday effect: in a magnetized plasma, the dielectric constant is a tensorial quantity, so components of the light with different polarizations propagate at different speeds. This magnetically induced birefringence produces a characteristic *Faraday rotation* of χ that is a function of the plasma properties along the line of sight

$$\Delta\chi = \frac{1}{2} \int \mathrm{d}l \, \rho_V \tag{2.8}$$

$$\approx \frac{1}{\nu^2} \frac{e^3}{2\pi m_e^2 c^2} \int \mathrm{d}l \, n_e f_{\mathrm{rel}}(\Theta_e) B_{||}, \tag{2.9}$$

where ρ_V is the Faraday rotation coefficient.[4] In the second line, we have substituted a high-frequency expression for ρ_V for the $<$ ultra-relativistic case; here, $B_{||}$ is the component of the magnetic field along the line of sight, and $f_{\mathrm{rel}} \approx \log \Theta_e / 2\Theta_e$ for the relativistic $\Theta_e \gg 1$ case and ≈ 1 otherwise (e.g., Rybicki & Lightman 1979; Quataert & Gruzinov 2000, see also Shcherbakov 2008 for a discussion). Evidently the observed polarization pattern is related

[3]We use the International Astronomical Union convention (see Hamaker & Bregman 1996) in which positive Q is oriented North–South (vertically) and positive U is oriented along the NorthEast–SouthWest direction (top left to bottom right). In this convention, EVPA is measured East of North, i.e., counter-clockwise from vertical on the sky.

[4]The factor of 1/2 is the same as in Equation 2.7. This can be seen by differentiating Equation 2.7 and substituting in the derivatives of Q and U from Equation 2.10 in the case of no emission, absorption, or circular polarization.

to the structure of the magnetic field in the accretion flow in a complicated way.

The full description of the polarized intensities can be evolved by solving the polarized radiative transfer equation along the length of the geodesic. In order to produce millimeter radio polarized images, we need only treat the emission and absorption processes, since scattering contributes negligibly at the relevant frequencies and plasma parameters. Expanding Equation 2.5 for emission, absorption, and mixing, the nonrelativistic, polarized radiative transfer equation is

$$\frac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix} I_\nu \\ Q_\nu \\ U_\nu \\ V_\nu \end{pmatrix} = \begin{pmatrix} j_{\nu,I} \\ j_{\nu,Q} \\ j_{\nu,U} \\ j_{\nu,V} \end{pmatrix} - \begin{pmatrix} \alpha_{\nu,I} & \alpha_{\nu,Q} & \alpha_{\nu,U} & \alpha_{\nu,V} \\ \alpha_{\nu,Q} & \alpha_{\nu,I} & \rho_{\nu,V} & -\rho_{\nu,U} \\ \alpha_{\nu,U} & -\rho_{\nu,V} & \alpha_{\nu,I} & \rho_{\nu,Q} \\ \alpha_{\nu,V} & \rho\nu,U & -\rho_{\nu,Q} & \alpha_{\nu,I} \end{pmatrix} \begin{pmatrix} I_\nu \\ Q_\nu \\ U_\nu \\ V_\nu \end{pmatrix}, \tag{2.10}$$

where the emissivities j_ν, absorptivities α_ν, and rotativities ρ_ν are frame-dependent quantities (see Landi Degl'Innocenti & Landi Degl'Innocenti 1985).

To generalize away from frame-dependence, Broderick & Blandford (2004) described a method to write the transport equation in terms of quantities that account for rotations of the observer frame along the line of sight. To provide a manifestly covariant description, Gammie & Leung (2012) produced an alternative formulation of the polarized transport equation in terms of a Hermitian coherency tensor $N^{\alpha\beta}$, which can be related to the Stokes parameters by the Pauli matrices. The polarized radiative transfer equation can then be written in a manifestly covariant form

$$k^\mu \nabla_\mu N^{\alpha\beta} = J^{\alpha\beta} + H^{\alpha\beta\gamma\delta} N_{\gamma\delta}, \tag{2.11}$$

where ∇_μ is the covariant derivative along the geodesic, $J^{\alpha\beta}$ is an emissivity tensor, and $H^{\alpha\beta\gamma\delta}$ is a tensor that accounts for absorption and rotation. Using this description to solve the polarized transport equation amounts to evaluating the nonrelativistic emissivities, absorp-

tivities, and rotativities in any convenient frame, constructing $J^{\alpha\beta}$ and $H^{\alpha\beta\gamma\delta}$, and solving Equation 2.11. In practice, it is often easiest to construct $J^{\alpha\beta}$ and $H^{\alpha\beta\gamma\delta}$ in the tetrad basis defined by the fluid four-velocity, the photon wavevector, and the orientation of the local magnetic field.

In some cases, we may solve an approximate form of the radiative transfer equation that does not account for polarization,

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{I_\nu}{\nu^3}\right) = \left(\frac{j_\nu}{\nu^2}\right) - (\nu\,\alpha_{\nu,I})\,\left(\frac{I_\nu}{\nu^3}\right). \tag{2.12}$$

(Here we have expressed all quantities in their relativistically invariant forms.) This approximation is sometimes called *unpolarized* transport even though it accounts for the total intensity of the light rather than only the unpolarized component.

Scattering

Photons undergo (inverse) Compton scattering as they travel through the plasma in the accretion flow. The inverse Compton scattering events allow the relativistic (high energy) electrons in the accretion flow to upscatter photons in the radiation field to higher energies. In our context, upscattering happens more frequently than downscattering. Collectively, these scattering events affect the spectrum of the accretion flow as observed at a large distance; this process is known as Comptonization. The importance of Comptonization can be gauged by the so-called Compton y parameter

$$y = N_{\mathrm{scatterings}} \times (\text{energy gained per scattering}) \tag{2.13}$$

$$\simeq \max\left(\tau_e, \tau_e^2\right) \times \left(4\Theta_e + 16\Theta_e^2\right), \tag{2.14}$$

where the optical depth to a scattering event within a system of size l is

$$\tau_e \approx n_e \sigma_T l. \tag{2.15}$$

In Equation 2.14, we have written the approximate expression for y that covers relativistic energies (see Zdziarski 1985).

The number density rate of interactions between two species of particles can be written

$$\dot{n}_{\rm int} = n_1 n_2 \left\langle \sigma_{12} v_{12} \right\rangle , \tag{2.16}$$

where n_1 and n_2 are the two species' number densities, σ_{12} is their interaction cross section, and v_{12} is their relative velocity. In flat space, the rate of interaction between a photon with wavevector k^μ and massive particles in a distribution function $\mathrm{d}n_m/\mathrm{d}^3p$ is

$$\dot{n}_{\rm int} = \int \mathrm{d}^3p \frac{\mathrm{d}n_m}{\mathrm{d}^3p} \frac{-k_\mu p^\mu}{k^t} \sigma_{m\gamma}\, c \tag{2.17}$$

$$= n_m \sigma_h\, c. \tag{2.18}$$

In the second line, we define a "hot cross section" (see Appendix III of Canfield et al. 1987)

$$\sigma_h \equiv \int \mathrm{d}^3p \frac{\mathrm{d}n}{\mathrm{d}^3p} \left(1 - \mu_m \beta_m\right) \sigma_{m\gamma}\, c, \tag{2.19}$$

where we have rewritten the dot product in terms of the particle speed in the plasma frame β_m and the cosine of the angle between the particle momentum and the photon momentum μ_m, both evaluated in a convenient frame.

Given an electron distribution function, the scattering process amounts to (1) probabilistically picking an electron from the distribution function according to the hot cross sections so that the incoming photon can scatter off of it and (2) probabilistically scattering the photon to a new wavevector (energy + direction) according to the Klein–Nishina cross section of the chosen electron. Here, the Klein–Nishina cross section is

$$\sigma_{\rm KN} = \sigma_T \frac{3}{4\epsilon^2} \left(2 + \frac{\epsilon^2 \left(1+\epsilon\right)}{\left(1+2\epsilon\right)^2} + \frac{\epsilon^2 - 2\epsilon - 2}{2\epsilon} \log\left(1+2\epsilon\right) \right), \tag{2.20}$$

where σ_T is the Thompson cross section and the energy of the photon in the electron rest frame is $\epsilon = -p_\mu k^\mu / m_e^2$.

2.2.3 Connecting GRMHD and GRRT

We produce the above-described image and spectrum electromagnetic observables from the fluid simulations (which output a description of the fluid state) by performing radiative transfer calculations. In order to use the GRMHD output, we must carefully treat several points.

Scaling and orienting GRMHD output

Unlike in GRMHD, the equations of radiative transfer are not scale invariant, so the numerical fluid data must be translated into physical units in order to perform the ray tracing. The use of GRRT introduces two scales: a length scale and a density scale. The length scale $\mathscr{L}$ determines the absolute size of the accretion system and is often written in terms of the mass of the central black hole $\mathscr{L} = GM/c^2$. The length scale sets the characteristic timescale, $\mathscr{T} = \mathscr{L}/c = GM/c^3$.

The density scale $\mathscr{M}$ provides units to the fluid rest-mass density, internal energy, and magnetic field strength while respecting the constraint that magnetization $\sigma = b^2/\rho$ and plasma β are independent of the choice of units, i.e., $\rho \to \mathscr{M}\rho$, $b^2 \to \mathscr{M}b^2$, and $u \to \mathscr{M}u$. The density scale is chosen so that the simulated images match two observational constraints: the image-integrated flux must be correct and the extent of the compact flux region must be reasonable. The latter constraint amounts to ensuring that the accretion flow is $\sim$ optically thin at 230 GHz. When possible, the resultant mass-accretion rate is compared to the predictions of single-zone models (see EHTC V; EHTC VIII) and observational estimates of $\dot{M}$ from rotation measure.

Finally, the accretion flow must be oriented properly with respect to the camera location on Earth. This amounts to choosing two of the three Euler angles: an inclination angle i

between the spin axis of the system and the elevation of the camera, and a position angle for the camera (the orientation that the projected spin axis makes on the image plane). The third angle is unnecessary because of the axisymmetry of the accretion flow. The inclination and position angle are free parameters in general, although large-scale information about the system, such as the orientation of a jet, may provide extra hints about how to proceed.

Assigning electron temperatures

Although our GRMHD simulations often make the approximation that the plasma is thermal and described by a single temperature, we account for the likely collisionless nature of the flow by allowing electron temperatures to deviate from the ion temperatures. When not using the entropy tracking procedure of Ressler et al. (2015), we assign electron temperatures according to the prescription of Mościbrodzka et al. (2016, see also Mościbrodzka et al. 2017; Event Horizon Telescope Collaboration et al. 2019e), wherein the ion-to-electron temperature ratio is determined by the local plasma β. This prescription is motivated by the idea that heating may be stronger in a strongly magnetized plasma (e.g., Quataert 1998; Quataert & Gruzinov 1999).

As in Mościbrodzka et al. (2016), the temperature ratio is parameterized by $R_{\rm low}$, $R_{\rm high}$, and $\beta_{\rm crit}$, with

$$R \equiv T_i/T_e = \frac{R_{\rm low} + R_{\rm high}\tilde{\beta}^2}{1+\tilde{\beta}^2}, \tag{2.21}$$

where $\tilde{\beta} \equiv \beta/\beta_{\rm crit}$.

To recover the electron temperature from the total fluid energy, we partition the fluid energy into two components associated with the electrons and with the ions.[5] Assuming an

[5]Some codes instead set the ion temperature equal to the fluid temperature, which over-counts the energy in the system.

ideal gas equation of state, the energies associated with the ions and electrons will be

$$u_i = (\hat{\gamma}_i - 1)^{-1} n_i k T_i \tag{2.22}$$

$$u_e = (\hat{\gamma}_e - 1)^{-1} n_e k T_e, \tag{2.23}$$

where $\hat{\gamma}_i$ and $\hat{\gamma}_e$ are the adiabatic indices of the ions and electrons respectively. The ion and electron number densities are related to the total mass density by

$$n_e = y\rho/m_p \tag{2.24}$$

$$n_i = z\rho/m_p, \tag{2.25}$$

where y and z are the number of electrons and nucleons per unionized atom, respectively.[6] The electron temperature is thus

$$T_e = \frac{m_p u \left(\hat{\gamma}_e - 1\right)\left(\hat{\gamma}_i - 1\right)}{k\rho \left(\left(\hat{\gamma}_i - 1\right) y + \left(\hat{\gamma}_e - 1\right) R z\right)}. \tag{2.26}$$

We typically assume that the ions in the plasma are nonrelativistic, so $\hat{\gamma}_i = 5/3$, and that the electrons are relativistic, so $\hat{\gamma}_e = 4/3$. Since we assign electron temperatures after the fact, we assume that each of $\hat{\gamma}_i, \hat{\gamma}_e$, and $\hat{\gamma}$ is both constant and fixed across the entire simulation domain. This treatment is not entirely self-consistent, since the adiabatic index of the total fluid should change depending on local contributions from the ion and electron fluid components (see Sądowski et al. 2017 for a self-consistent treatment).

Approximations and pathologies

GRMHD schemes are not robust in regions with high magnetization $\sigma \gg 1$ or where plasma $\beta \ll 1$. In order to ensure numerical stability, σ and β are computed in each cell of the simulation domain for each time step, and mass or internal energy is injected into simulation

[6] Notice that y and z are not the conventional Y and Z mass fractions.

zones to ensure that neither σ and $1/\beta$ exceed some preset ceilings. The values of the ceilings are typically varied across different trial-run simulations to ensure that they have minimal impact on the dynamical evolution of the flow.

In black hole accretion simulations, the ceilings are generally triggered in the highly magnetized jet funnel regions near the poles, where we expect $\rho \ll 1$. Although we do not trust the fluid thermodynamics in regions where the ceilings are activated, we do expect that the true plasma densities in the jet region will be low enough that no appreciable emission is produced. We therefore zero all emission that would be produced in any regions where $\sigma > 1$. We also zero emission within a few degrees of the poles, since some treatments of the polar boundary condition may cause σ to artificially drop below 1 there. The effect of the σ cutoff is considered in the detailed analyses (see, e.g., Chael et al. 2019; EHTC VIII) and briefly in § 2.6.

2.3 Code detail

We now describe the `PATOKA` pipeline (see Figure 2.1), in which simulated observables are generated by ray tracing snapshots produced from numerical GRMHD simulations of the black hole accretion flows. These two stages, GRMHD and GRRT, are separated for computational efficiency, since GRMHD simulations are costly and multiple radiation models can be applied to a single fluid snapshot without rerunning the fluid simulation. We now describe the details of the three codes we use and place a particular emphasis on differences between the versions we use and the code as described upon release. All of the codes used in the `PATOKA` pipeline compute metric derivatives numerically and therefore support arbitrary metrics through redefinition of the metric line element.

Figure 2.1: An oil field in Marion County, Illinois. `PATOKA` is named in reference to the Patoka Oil Terminal, which serves major pipelines in the second district of the Petroleum Administration for Defense Districts. Image credit: Arthur Rothstein / Farm Security Administration Office Of War Information.

2.3.1 General relativistic magnetohydrodynamics

We use the `iharm` code (Gammie et al. 2003; Noble et al. 2006, 2009) to integrate the equations of ideal GRMHD. Porth et al. (2019) provides a comparison of contemporary GRMHD codes in the context of SANE accretion flows. `iharm` is a conservative second-order explicit shock-capturing finite-volume code for arbitrary stationary spacetimes, a descendant of `harm2d` (J. Dolence, priv. communication), and based on the `harm` scheme of Gammie et al. (2003). The code is publicly available.[7] Validation and scaling tests are described in Gammie et al. (2003) and show second-order convergence on a suite of test problems.

The governing equations of ideal GRMHD take the form of a system of hyperbolic partial differential equations. In conservative form and written in a coordinate basis, the equations

[7] `https://github.com/afd-illinois/iharm3d`

are

$$\partial_t \left(\sqrt{-g}\rho u^t\right) = -\partial_i \left(\sqrt{-g}\rho u^i\right), \tag{2.27}$$

$$\partial_t \left(\sqrt{-g}T^t{}_\nu\right) = -\partial_i \left(\sqrt{-g}T^i{}_\nu\right) + \sqrt{-g}T^\kappa{}_\lambda \Gamma^\lambda{}_{\nu\kappa}, \tag{2.28}$$

$$\partial_t \left(\sqrt{-g}B^i\right) = -\partial_j \left[\sqrt{-g}\left(b^j u^i - b^i u^j\right)\right], \tag{2.29}$$

$$\partial_i \left(\sqrt{-g}B^i\right) = 0, \tag{2.30}$$

where the plasma is defined by its rest mass density ρ_0, its four-velocity u^μ, and b^μ is the magnetic field four-vector following McKinney & Gammie (2004). Here, $g \equiv \det(g_{\mu\nu})$ is the determinant of the covariant metric, Γ is a Christoffel symbol, and i and j denote spatial coordinates. In Equations 2.29 and 2.30, we leverage the ideal MHD condition[8] $u_\mu F^{\mu\nu} = 0$ to express the electromagnetic field tensor $F^{\mu\nu}$ in terms of $B^i \equiv {}^\star F^{it}$ for notational simplicity. The stress–energy tensor $T^\mu{}_\nu$ contains contributions from both the fluid and the electromagnetic field

$$T^\mu{}_\nu = \left(\rho + u + P + b^\lambda b_\lambda\right) u^\mu u_\nu + \left(P + \frac{b^\lambda b_\lambda}{2}\right) g^\mu{}_\nu - b^\mu b_\nu, \tag{2.31}$$

where u is the internal energy of the fluid and the fluid pressure P is related to its internal energy through a constant adiabatic index $\hat{\gamma}$ with $P = (\hat{\gamma} - 1)\, u$.

The equations are solved over a logically Cartesian, three-dimensional grid in arbitrary coordinates. Eight primitive variables,[9] ρ, u, V^i, and B^i, are stored at the center of each grid zone and evolved forward across Kerr–Schild time steps. Fluxes are computed using the local Lax–Friedrichs method (Rusanov, 1962), and the divergence constraint of Equation 2.30 is enforced using the Flux–CT scheme described in Tóth (2000, see also Balsara & Spicer 1999). The fifth-order WENO5 scheme (Jiang & Shu, 1996) is typically used for spatial reconstruction. The code is parallelized across contiguous domain chunks using MPI (Forum,

[8] In general, the electric field in the frame of the plasma is $E^\mu = u_\nu F^{\mu\nu}$. In ideal MHD, the infinite conductivity of the plasma allows charges to reposition themselves quickly enough so that $E^\mu = 0$.

[9] Eight variables are saved for basic, ideal GRMHD. Extra variables may be tracked in extensions.

1994) and within each chunk using OpenMP (Dagum & Menon, 1998). `iharm` imposes a constant floor on plasma β and a geometric floor on ρ. It also imposes ceilings on σ, Θ_e, the fluid velocity measured with respect to the coordinate frame, and optionally the fluid entropy $P_{\rm gas}/\rho^{\hat{\gamma}}$.

Simulations are typically run in augmented versions of the horizon-penetrating modified Kerr–Schild (MKS) coordinates introduced in Gammie et al. (2003). In MKS, the three spatial coordinates x^1, x^2, x^3 are direct functions of radius, latitude, and azimuth respectively. The inner radial boundary of the simulation is chosen to ensure that $\gtrsim 5$ zones lie within the event horizon. The outer edge of the boundary is chosen so that the torus lies comfortably within the simulation domain. We use outflow boundary conditions along the two radial boundaries, a periodic boundary condition in the azimuthal direction, and a pseudo-reflecting boundary condition at the two latitudinal poles that mirrors the x^2 elevational components of the magnetic field and fluid velocity across the one-dimensional border.

`iharm` supports the subgrid electron heating method of Ressler et al. (2015) to track numerical dissipation and consistently heat electrons with some fraction of the total dissipated energy according to a prescription that is often motivated by kinetic studies of magnetic reconnection, turbulent heating, shock heating, or Ohmic heating. When the electron thermodynamics module is active, the code outputs two extra variables—functions of the total fluid entropy and the electron entropy—in addition to the base eight primitive variables listed above.

We initialize the fluid sector of our accretion disk simulations with a Fishbone & Moncrief (1976, hereafter FM) torus. We seed the initial condition with a small random perturbation in u away from equilibrium in order to set off the magneto-rotational instability (Balbus & Hawley, 1991) and begin accretion onto the hole.[10] We parameterize the FM torus by its inner edge at $r = r_{\rm in}$ and the radius of the maximum pressure at $r = r_{\rm max}$.

[10] An unperturbed torus is also unstable to the Papaloizou & Pringle (1984) instability.

We simulate both SANE and MAD accretion flows. Although the steady state magnetic flux is not trivially related to the initial conditions of the simulation, we implement configurations that have been identified in previous work and shown to produce either SANE and MAD flows. Mechanically, the initial condition for the magnetic field is determined from a prescribed, axisymmetric electromagnetic vector potential $A_\phi(r, \phi)$, which is computed at simulation zone corners.

SANE disks are generated from

$$A_\phi = \max\left[\frac{\rho}{\rho_{\max}} - 0.2, 0\right], \tag{2.32}$$

where $\rho_{\max}$ is the maximum initial plasma density. MAD disks are generated from

$$A_\phi = \max\left[\frac{\rho}{\rho_{\max}}\left(\frac{r}{r_0}\sin\theta\right)^3 e^{-r/400} - 0.2, 0\right], \tag{2.33}$$

where r_0 is chosen to be the inner boundary of the simulation domain. This prescription concentrates the initial magnetic field toward the inner edge of the disk and forces it to taper at large r. (See Figure 5.1 and the surrounding discussion for examples.)

2.3.2 Ray-tracing images

`ipole` (Mościbrodzka & Gammie, 2018) is a publicly available GRRT code for covariant polarized radiative transfer and is a descendant of the unpolarized image code `ibothros` (Noble et al., 2007). `ipole` produces a rectangular polarimetric image defined by a field of view (width in GM/c^2, or translated to μas in the context of EHT sources), distance from the black hole ($d_{\rm source}$), and orientation with respect to the black hole spin axis and midplane (inclination and position angle). Each pixel reports the specific intensities for the Stokes parameters $I_\nu, Q_\nu, U_\nu, V_\nu$ at the pixel center as well as the total optical and Faraday depths along the pixel-centered geodesic.

`ipole` is an observer-to-emitter code, meaning that only geodesics that intersect a pre-

defined pinhole camera (observer) are considered. The camera is defined by a particular coordinate x^μ_c and an orthonormal tetrad specified by the normal observer velocity,[11] a radially directed wavevector, and the black hole spin axis. The pinhole camera defines an image with pixels that all intersect x^μ_c and are regularly spaced in angle as measured in the tetrad frame. This prescription ensures that the central pixel corresponds to a geodesic with impact parameter zero. The geodesic for each pixel is integrated backwards toward the black hole (i.e., the emitting matter) until it leaves the simulation domain. The radiative transfer equation is then solved forwards toward the camera.

`ipole` solves Equation 2.11 in two independent stages. In the first stage, it numerically integrates and parallel-transports the coherency tensor $N^{\alpha\beta}$ along the geodesic. In the second stage, it projects $N^{\alpha\beta}$ into a Stokes vector $S = (I, Q, U, V)$ in a local fluid tetrad,[12] evaluates transfer coefficients in that tetrad, and uses an analytic solution to update S before translating it back to $N^{\alpha\beta}$ for the next step. The analytic solution ensures numerical stability when the plasma has large optical or Faraday depths, since the geodesic step size is not determined by the local plasma properties. The above approach is coordinate agnostic, since all geometric factors are computed numerically from the metric. A comparison of contemporary unpolarized GRRT codes is available in Gold et al. (2020).

The Illinois version of `ipole`, which is used in the `PATOKA` pipeline, implements several new features and differs from the originally published version in several ways, which we describe below. Both the originally published version of `ipole` and the Illinois version are publicly available.[13] `ipole` converges at second order.

[11] Earlier camera prescriptions defined the camera in the frame with $u^\mu \propto (1, 0, 0, 0)$.

[12] Because we treat synchrotron radiation, it is convenient to construct the tetrad from the fluid u^μ, the geodesic wavevector k^μ, and the local magnetic field orientation b^μ.

[13] The version of `ipole` described in this document can be found at `https://github.com/afd-illinois/ipole` as version 1.4.

Tracking the emission source

Since we aim to connect observables to the underlying fluid and black hole physics, it is useful to be able to study where the emission is produced. In order to compute what fraction of the emission produced by a given volume will contribute to an image, it is necessary to specify the location of the observer, since the emission coefficients depend on the photon wavevectors k^μ through the local sampling frequency $\nu = -k^\mu u_\mu$ and the pitch angle $= \arccos(k_\mu b^\mu)$. Furthermore, strong lensing allows multiple geodesics to sample the same region of space; since each geodesic has its own k^μ, a bijective mapping between points in space and emissivities does not exist.

Moreover, not all emitted light makes it to the observer. Light emitted near the event horizon may fall into the hole, and the optical depth of the plasma between emission and the observer will cause some light to be absorbed. It is straightforward to compute what fraction of light makes it to the observer at infinity for the Stokes I total intensity by saving the values computed at each geodesic step when solving the approximate radiative transfer equation Equation 2.12.

We save the local optical depth and the local contribution to the intensity $\mathrm{d}I$ at each step when solving the radiative transfer equation along the geodesic. After the full geodesic has been traced, the absorptivity to any point along the geodesic can be computed as the exponential of the total optical depth. The product of the local intensity contribution with the absorption is computed per geodesic step and saved in an array representing the simulation domain. The process is validated after the complete image has been generated by summing the flux contributed by each zone and comparing to the final total image flux.

Subring decomposition

Synthetic images of RIAF GRMHD simulations display a clear ring-like structure. This structure is identified with the photon ring produced as strong lensing allows light paths to wind around the black hole multiple times (see, e.g., Johnson et al. 2020). In aggregate,

the composition of the subrings produces a quasi-exponentially divergence in brightness temperature that is limited in part by the optical depth of the plasma.

Observable signatures of the photon ring structure have been probed by analytic work (e.g., Gralla & Lupsasca 2020; Himwich et al. 2020; Vincent et al. 2021), but the detailed structure of the observable photon ring may be heavily influenced by the structure of the emitting plasma. We have included a subring decomposition routine in `ipole` that allows the code to produce separate images for each subring. The subring structure is particularly evident when the emission exhibits a non-trivial latitudinal structure, e.g., if the emission is concentrated near the midplane and each subring contribution can be easily separated from its neighbors.

Geodesics that pass close to the hole experience strong lensing and may undergo latitudinal oscillations—this effect is most obvious for bound orbits, which circle the hole indefinitely and undergo periodic oscillations as they sweep through a range of latitudes. Each event x^μ along a geodesic can be assigned an orbit-number n, corresponding to the number of latitudinal turning points $\theta_\pm$ between the event and the observer. This value can be directly tracked and saved during the backward geodesic integration.

In order to synthesize the $n^{\rm th}_{\rm target}$ subring image, the emission coefficients $j_{I,Q,U,V}$ are zeroed during the forward radiative transfer integration in regions of the geodesic with $n \neq n_{\rm target}$. It is important to include parallel transport as well as absorption and rotation in regions with $n \neq n_{\rm target}$, since although the subring image comprises only photons that were emitted along the corresponding geodesic segment, we must account for how the photons interact with matter and spacetime as they propagate to the observer.

Differences from the original version

We have slightly modified `ipole` from the originally published version in several ways:

1. `ipole` has been modified to read fluid snapshot files with arbitrary logical coordinate systems through the use of the `simcoords` module. Using `simcoords`, `ipole` performs

all ray-tracing in exponential Kerr–Schild (eKS) coordinates and uses an interpolated grid map between eKS coordinates and the input snapshot coordinates. In order to make use of the `simcoords` module, a fluid snapshot need only describe the location of each grid zone in Kerr–Schild coordinates and provide the velocity and magnetic-field vectors with Kerr–Schild components.

2. Several different methods for treating the σ magnetization cutoff have been implemented. The default version treats σ as an interpolated scalar and simply zeros the transfer coefficients in regions with $\sigma > \sigma_{\rm cutoff}$ by setting the number density of electrons $n_e = 0$. Alternative methods include gradually suppressing n_e according to a sigmoidal function of σ or zeroing n_e on a per-zone basis according to zone-centered values of σ. The choice of cutoff procedure may introduce image artifacts—see § 2.6 for more detail.

3. When evaluating fluid parameters along geodesics, we interpolate the scalars n_e, u/ρ, magnetic field strength as well as the six primitive variables that describe the fluid velocity and magnetic field orientation. This procedure ensures that temperatures, which are derived from u/ρ remain reasonable and that the interpolation scheme does not lower the magnetic field strength in regions where the magnetic field oscillates wildly, like near the jet–disk boundary. Then, by constructing u^μ and b^μ from the primitives, we can ensure $u^\mu u_\mu = -1$ and $u^\mu b_\mu = 0$. We have observed that interpolating the four-vector components directly tends to produce $\geq \mathcal{O}(1)$ deviations from these criteria within $\approx 2\ GM/c^2$ of the event horizon. Deviations have been observed to be catastrophic in some cases when the `simcoords` module is used.

2.3.3 Producing spectra

`igrmonty` (Dolence et al., 2009) is a Monte Carlo radiative transport code designed to generate spectra from GRMHD fluid simulation snapshot files of optically thin ionized plasmas.

It accounts for the full angle- and frequency-dependence of emission and absorption, and it treats single Compton scattering exactly. `igrmonty` is an emitter-to-observer code, meaning it simulates emission at all frequencies and angles across the entire domain. This procedure is slower than the observer-to-emitter procedure, but it more cleanly and easily accounts for probabilistic scattering into the light of sight. The observer-agnostic treatment naturally produces the full spectrum νL_ν as a function of elevation and azimuth around the black hole.

`igrmonty` tracks a Monte Carlo sample of the radiation field in the form of *superphotons*. A superphoton with weight w is a packet of $w \gg 1$ photons, where each photon has the same position x^μ and wavevector k^μ. Superphotons are created across the computational domain according to the local emissivity of the plasma. In order to emit a target number $N_{\rm target}$ of superphotons over the full simulation domain, the bolometric luminosity due to emission is precomputed for each zone and used to determine what fraction of the $N_{\rm target}$ total superphotons should be emitted per zone. This heuristic can produce a noisy signal when the fluid simulation resolution is increased if there is non-trivial structure in the accretion flow, thus higher fluid resolutions typically require a larger number of superphotons during the spectrum-generation step.

To optimize signal in the spectrum, the weight factor w is conventionally chosen per frequency according to the heuristic that each logarithmic bin in energy space should contain approximately the same number of superphotons—note that we deviate slightly from this procedure when including bremsstrahlung emission. Since it is impossible to predetermine how many superphotons survive extinction on their journeys to the observer, we estimate how the pre-computed weight factors should be set by assuming that all emitted photons escape to infinity. We also neglect factors like redshift, scattering, and the ultimate angular dependence of the spectrum. When a new superphoton is created, its frequency is chosen according to rejection sampling and the rest of its wavevector is initialized in a local orthonormal tetrad according to the pitch-angle-dependent emissivity prescription. Each superphoton saves its initial position; this information can be used to infer properties of the emission region.

After emission, the optical depths to both absorption and scattering are computed as the superphotons propagate along their geodesics. Absorption is accounted for by decreasing the weight w of the superphoton packet, which is directly proportional to the intensity of the radiation packet.

`igrmonty` treats scattering in a probabilistic sense: a superphoton will scatter with some probability at each step along its geodesic; if it scatters, the wavevector of the scattered superphoton is evaluated in a local orthonormal tetrad and determined probabilistically from the differential electron scattering cross section. In order to boost the signal in the Compton-upscattered component of the spectrum, `igrmonty` uses a variant of the bias method introduced in Kahn (1950). Although the likelihood p that a full superphoton will scatter may be small in the optically thin plasmas we consider, the bias procedure enables a fraction $1/b$ of a fiducial superphoton to scatter with probability bp. If the superphoton scatters, its weight is decreased to $1 - 1/b$ and a new superphoton representing the scattered component is created with weight $1/b$.

`igrmonty` is publicly available,[14] converges at second order in geodesic integration, and converges like $\sqrt{N_s}$ in the number of Monte Carlo radiation field samples. The `PATOKA` version of `igrmonty` only solves the approximate (unpolarized) radiative transfer equation, but recently Mościbrodzka (2020) introduced the `RADPOL` scheme, which accounts for fully polarized synchrotron emission, absorption, Faraday rotation and conversion, and Compton scattering.

Scattering bias factor

It is difficult to determine how the bias should be set before running the simulation. If it is too low, too few superphotons will scatter and the Compton contribution to the spectrum will be unusably noisy. If it is too high, the code could reach a "supercritical" state, in which superphotons produced through scattering also undergo scattering events and the

[14] `https://github.com/afd-illinois/igrmonty`

total number of superphotons diverges. Ideally, the bias should be set so that the number of superphotons created through scattering is commensurate with the number emitted. In all cases, the bias factor must be ≥ 1.

The bias parameter is typically scaled per zone by a trial bias factor that is set by the square of the local electron temperature Θ_e^2 to improve resolution in the scattering events with higher energies compared to the average. This trial bias is multiplied by a global tuning factor $b_{\rm tuning}$, which scales the scattering rate across the entire simulation. `igrmonty` begins each run with a low resolution "bias tuning" step, during which $b_{\rm tuning}$ is varied until the ratio between the number of superphotons created through scattering and the number created through emission is approximately unity. The ratio is tracked *in situ* during the tuning runs, and the evaluation is halted if the ratio exceeds a large number to ensure the supercritical state can be preempted.

Bremsstrahlung

Yarza et al. (2020) found that bremsstrahlung emission may dominate the high-energy ($\gtrsim$ 512 keV) component of the spectrum in radiative simulations of SANE accretion flows as $\dot{M}$ is increased. In relativistic plasmas, both electron–electron and electron–ion bremsstrahlung contribute to the total emissivity; `igrmonty` supports several contemporary prescriptions for both emissivities (see the Appendix of Yarza et al. 2020 for a comparison of the prescriptions). Since synchrotron and bremsstrahlung emission may be simultaneously non-zero, each emitted superphoton is assigned a bremsstrahlung fraction value between zero and one. This value is saved when the spectrum is recorded and can be used to determine how much emission should be associated with which emission process.

Bremsstrahlung emission dominates direct synchrotron at higher frequencies where the optical depth of the plasma is low. A naive application of the superphoton weight assignment scheme described above may thus preferentially improve signal in the bremsstrahlung component of the spectrum. This can be addressed by either: modifying the weighting proce-

dure to independently generate weights for the lower frequency synchrotron emission and the higher frequency bremsstrahlung emission; or independently generating spectrum from the synchrotron- and bremsstrahlung-emission components and adding them together. Compton scattering is typically unimportant for bremsstrahlung emission, so the bremsstrahlung calculation may be performed with scattering turned off.

Arbitrary electron distribution functions

`igrmonty` relies on the full electron distribution function both when computing the Compton scattering cross section and during the scattering procedure when choosing an electron of which to scatter. The version of the code published in Dolence et al. (2009) implemented a semi-analytic treatment of the cross section and scattering sampling routines; here, we instead use an analytic expression only for the distribution function itself and numerically evaluate the cross section. We use the rejection sampling technique to sample the distribution rather than rely on an inversion of the distribution function. Because this procedure relies only on a prescription for the distribution function, it is quite general and supports any isotropic distribution that can be written down as a function of local fluid parameters.

2.4 Sample data products

We now present example data products from each part of the `PATOKA` pipeline. The data we show here were generated for the EHT M87 simulation library that was used for validation in EHTC IV; EHTC VII and for analysis in EHTC V; EHTC VI; EHTC VIII.

2.4.1 GRMHD models

The GRMHD library generated by `PATOKA` for the EHT M87 analysis included two parts. The first part comprised ten "canonical" simulations spanning the MAD and SANE accretion

states and five spins $a_* = -15/16, -1/2, 0, +1/2, +15/16$.[15] Hereafter, we write 1/2 and 15/16 as 0.5 and 0.94 to be consistent with the EHT paper sequence. The second part was composed of several dozen ancillary simulations used to the initial condition, disk size, adiabatic index, and simulation resolution. Each of the models generated for the library was evolved until at least $t = 10^4\ GM/c^3$, during which time its accretion flow is believed to reach a statistical steady state within $r \leq 10 - 20\ GM/c^2$. GRMHD fluid snapshots were saved every $5\ GM/c^3$, corresponding to ≈ 43 hours for a black hole with mass $\approx 6.2 \times 10^9\ M_\odot$.

We only produced simulations with purely corotating (aligned) or counter-rotating (anti-aligned) accretion flows, since varying disk tilt adds another dimension to the parameter space and is thus prohibitively expensive. We only produced SANE simulations with $\phi \approx 1$, although initial conditions that produce both $\phi \ll 1$ and $1 \lesssim \phi \lesssim \phi_c$ accretion states are known. The fluid was assumed to have a uniform constant adiabatic index for each simulation $\hat{\gamma}$, although the value of $\hat{\gamma}$ was varied between different simulations.

Each simulation was run on a three-dimensional regular grid defined in the horizon-penetrating FMKS coordinates, which concentrate resolution toward the midplane and away from the jet at small radius (see § 2.7). We found that evolving a large disk over a long time could cause an initially SANE accretion flow to go MAD, so we chose to initialize the canonical SANEs with smaller accretion disks, allowing us to use smaller simulation domains and lower absolute resolutions.

Figure 2.2 shows snapshots of the plasma rest-mass density ρ and magnetization σ over the course of evolution of an intermediate-spin $a_* = 0.5$ MAD simulation. The initial condition with a large Fishbone–Moncrief torus and an ordered, looping magnetic field, progressively gives way to a turbulent accretion flow with a low-density high-magnetization funnel containing a strong, ordered magnetic field. The qualitative difference between the MAD and SANE accretion morphologies is shown in Figure 2.3. The MAD simulation is punctuated by magnetic bubbles corresponding to flux ejection events.

[15] Here, negative values of spin imply that the angular momentum of the hole and the disk are anti-parallel, i.e., a "retrograde" disk.

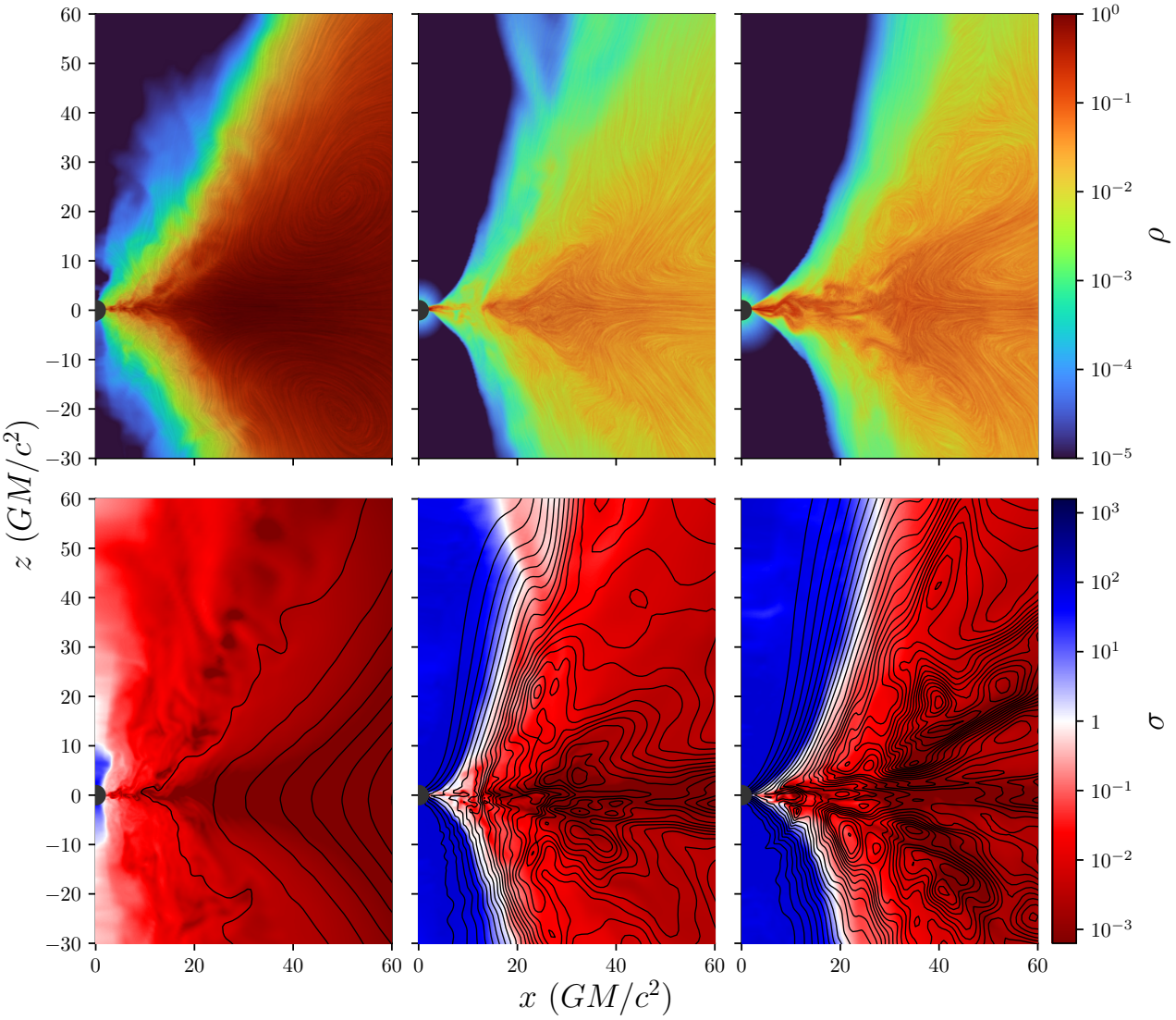


Figure 2.2: GRMHD simulation of a MAD accretion flow at $a_* = 0.5$. Top panels show rest-frame plasma density at three different times over the course of the simulation; the line integral convolution technique is used to represent the motion of the plasma. The bottom panels show magnetization with overplotted field lines at the same times as the top panels. The black hole event horizon is plotted as a dark circle. As the simulation evolves, the flow becomes increasingly turbulent and a high-magnetization jet region opens around the poles.

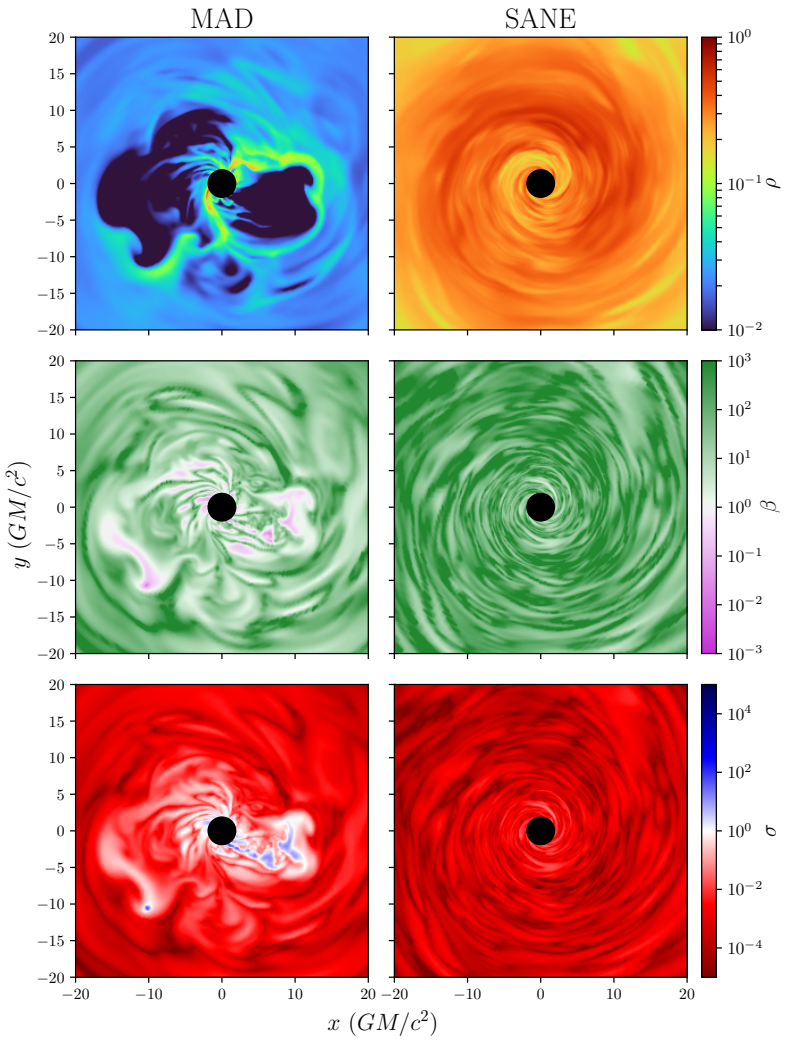


Figure 2.3: Comparison of density, plasma β, and magnetization σ in the midplane of MAD and SANE simulations with intermediate spin $a_* = 0.5$. In the MAD simulation, the accretion proceeds in thin strands, contrasting the steady but turbulent disk-like SANE accretion mode. The evacuated regions in the MAD simulation with low β and high σ are magnetic bubbles produced during the flux ejection events that episodically recur when an excess of magnetic flux has been trapped on the event horizon.

2.4.2 Electromagnetic observables

The GRMHD simulations were post-processed to generate images and spectra using `ipole` and `igrmonty`. In total, over 120000 images were generated for the canonical M87 total intensity and polarization analyses, and over 3 million were generated for supporting analyses, including resolution and field-of-view studies, explorations of the analyses robustness to changes in numerical parameters like geodesic step size, and machine learning projects. The full set of radiative transport model parameters is described below.

For the images produced for the EHT M87 sequence and discussed here, we assume that the accreting plasma is composed of pure ionized hydrogen, so that $y = z = 1$ in Equation 2.26. We fixed $\beta_{\rm crit} = 1$ and allowed the two parameters in Equation 2.21 to vary between $r_{\rm low} = 1, 10$ and $r_{\rm high} = 1, 10, 20, 40, 80, 160$. See Event Horizon Telescope Collaboration et al. (2019e) and Event Horizon Telescope Collaboration et al. (2021b) for a discussion of the motivation behind these choices.

The library discussed here was generated to compare against observations of M87, so it was generated using physical parameters that would target that system. The inclination angle i was chosen to be consistent with the orientation of the M87 jet at large scales, $i \approx 17^\circ$ (Hada et al., 2017; Kim et al., 2018; Walker et al., 2018), so we produced a library with inclinations ranging over multiple inclinations $i = 7, 12, 17, 22, 27$ degrees relative to the line of sight. We do not know *a priori* whether the black hole spin axis is directed toward or away from us. An exploratory survey of the library showed that it was necessary to orient the black hole spin vector away from Earth in order to reproduce both the image brightness asymmetry and the position angle of the large scale jet (Event Horizon Telescope Collaboration et al., 2019e). The position angle of the spin axis can be reoriented during analysis by rotating both the image and the per-pixel EVPA.

In order to perform radiative transfer, the GRMHD simulation mass density was scaled to physical units by requiring that the simulated compact flux at 230 GHz was consistent with a contemporaneous target flux drawn from observations (see Event Horizon Telescope

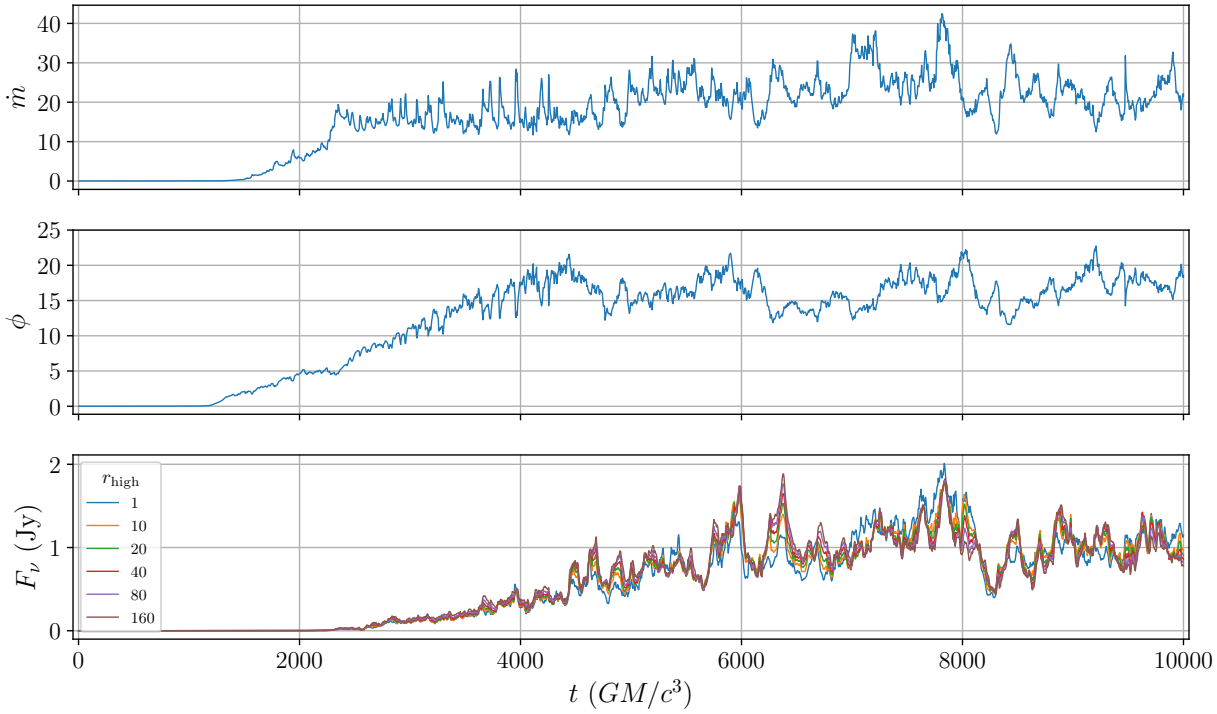


Figure 2.4: Time series of flux variables from the MAD $a_* = 0.5$ simulation. Top panel: mass accretion rate in arbitrary code units versus time. Center panel: dimensionless magnetic flux ϕ versus time. Bottom panel: Light curves at $\nu = 230$ GHz for $r_{\rm low} = 1$ and different $r_{\rm high}$ values. The light curves have been scaled to match a 1 Jy target flux over the last 5000 GM/c^3 of the simulation. The first half of the simulation is dominated by a transient from the torus initial condition. Notice that stability in $\dot{m}$ and ϕ does not necessarily equate to stability in the 1.3 mm flux.

Collaboration et al. 2019d for more detail on identifying the target flux). The relationship between the scaling factor $\mathscr{M}$ and total flux $F_{\rm tot}$ is a complicated function of the details of the accretion flow, but it tends to be monotonic near the target value, so identifying the appropriate scaling factor corresponds to a simple root-finding procedure. Since the fitting procedure is expensive, the flux is typically fit using the approximate total intensity solution over a regular subsample of the snapshots at low resolution. The quality of the fit is substantiated when the high resolution data are generated. After identifying the value of $\mathscr{M}$ required to produce the target flux, every snapshot from each GRMHD model is typically imaged, producing a sequence with a 5 GM/c^3 cadence.

The result of running the flux-fitting procedure is shown in Figure 2.4 for the canonical

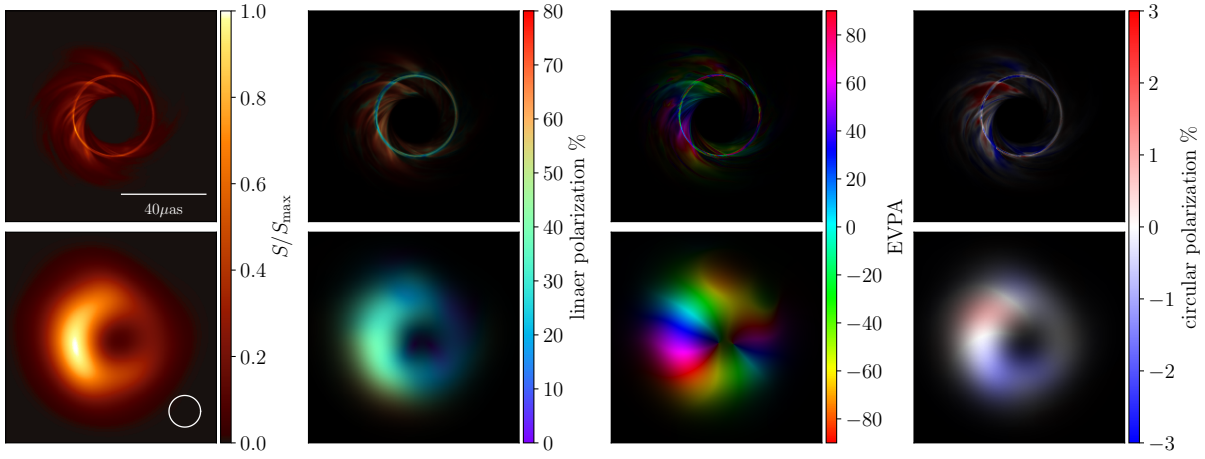


Figure 2.5: Example polarimetric images of MAD $a_* = 0.5$ simulation produced by `ipole` with (bottom) and without (top) blurring to 15 μas Gaussian beam. Panels show total intensity, linear polarization fraction, electric vector position angle (EVPA), and circular polarization fraction. Blurring decreases the observed linear polarization fraction in regions where EVPA is rapidly varying.

MAD $a_* = 0.5$ GRMHD simulation with $r_{\rm low} = 1$ and $i = 17°$. The unscaled mass accretion rate of the system $\dot{m}$ and the dimensionless magnetic flux parameter ϕ are also plotted for comparison. Here, the light curves have been fit so that the average 230 GHz flux matches 1 Jy over the last 5000 GM/c^3 of the simulation.

Figure 2.5 shows an example synthetic image from the MAD $a_* = 0.5$ model with $r_{\rm low} = 1$ and $r_{\rm high} = 40$ at $i = 17°$. Each pixel contains the full polarimetric Stokes I, Q, U, V specific intensities, which can be processed to provide information about the linear polarization fraction LP $= \sqrt{Q^2 + U^2}/I$, EVPA χ, and circular polarization fraction CP $= V/I$. Blurring is performed by convolving the Stokes intensities with a Gaussian beam.

Figure 2.6 shows the result of the ring decomposition procedure used to isolate the different subrings in a ray-traced image (see, e.g., Figure 3 of Johnson et al. 2020 for cross sections of a similar decomposition). Each next subring is exponentially demagnified compared to the last one; the image resolution and increasing optical depth of the source limit how many subrings contribute to the composite image.

`ipole` can be used to track the source of the observed flux, as seen in Figure 2.7. All

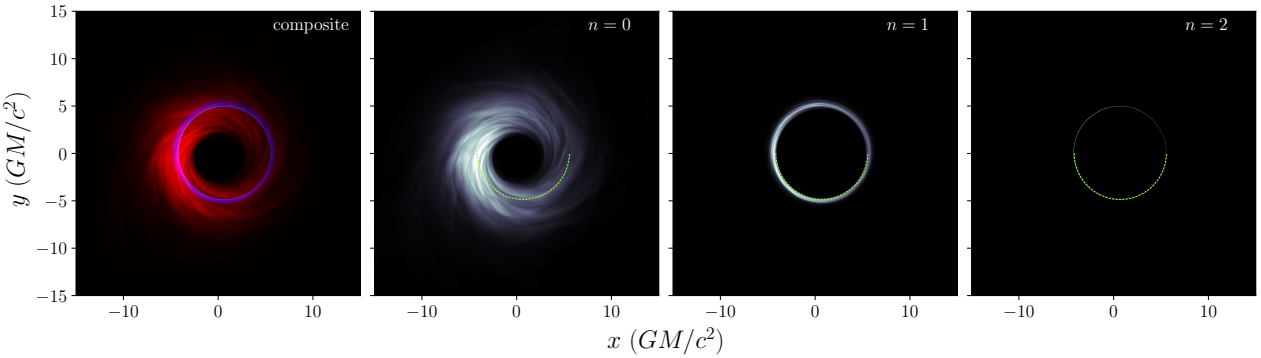


Figure 2.6: Decomposition of the total intensity into first three subrings for a snapshot from the high spin $a_* = 0.94$ MAD model with $r_{\rm low} = 1$, $r_{\rm high} = 10$, and $i = 17°$. The intensity produced in the $n = 0$, $n = 1$, and $n = 2$ subrings are shown in the right three panels, and the location of the critical curve is plotted as a dashed line in green. The left panel shows a composite image produced by adding the $n = 0$ subring in red, $n = 1$ subring in blue, and $n = 2$ subring in green.

MAD simulations tend to show the same characteristic fragmentary emission structure, which corresponds to the disjoint accretion (see, e.g., the top left panel of Figure 2.3). Much of the emission in the MAD case is thus produced in the hot, chaotic region of the flow near the horizon. SANE emission is comparatively more structured. Changing the electron temperature model in the SANE simulations can have drastic effects by heating up the jet funnel and shifting emission out of the disk.

Running `igrmonty` is significantly more computationally expensive than running `ipole`, so it is infeasible to generate spectra for every fluid snapshot across every radiation model. The two spectrum constraints considered in the EHT analysis were the overall radiative efficiency of the flow and a boolean determination if the simulation X-ray flux was consistently too high. At minimum, we generate ten spectra per radiation model, but we checked that producing a denser sampling of spectra in time does not change the statistical result. Figure 2.8 shows example spectra produced from one MAD and one SANE snapshot at $r_{\rm low} = 1$ and across different values of $r_{\rm high}$.

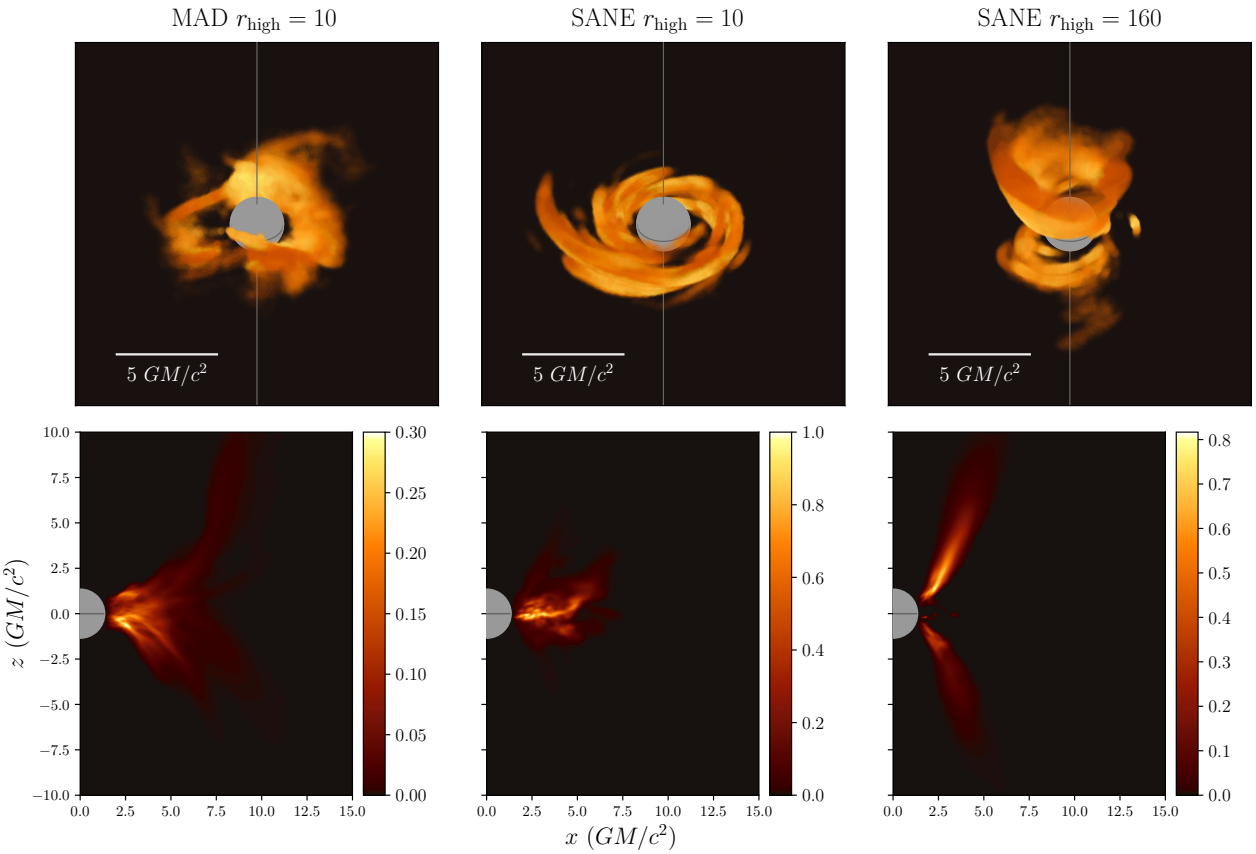


Figure 2.7: Location of emission around high spin black holes with $a_* = 0.94$. Top: three-dimensional render of emission source with color and transparency determined by the total emission produced within that region of space. Bottom: the same data as above after summing across the azimuthal dimension ϕ. The left column is a typical representation of emission from a MAD simulation; emission tracks the fragmentary plasma. The right two columns represent emission from the same SANE simulation snapshot but with different electron temperature prescriptions. Larger values of $r_{\rm high}$ shift the emission from the turbulent but steady disk into the funnel region. All simulations have $r_{\rm low} = 1$ and are imaged at $i = 17°$. The total flux produced by each simulation is the same, so the color scales show the relative concentration of emission in the azimuthal sum.

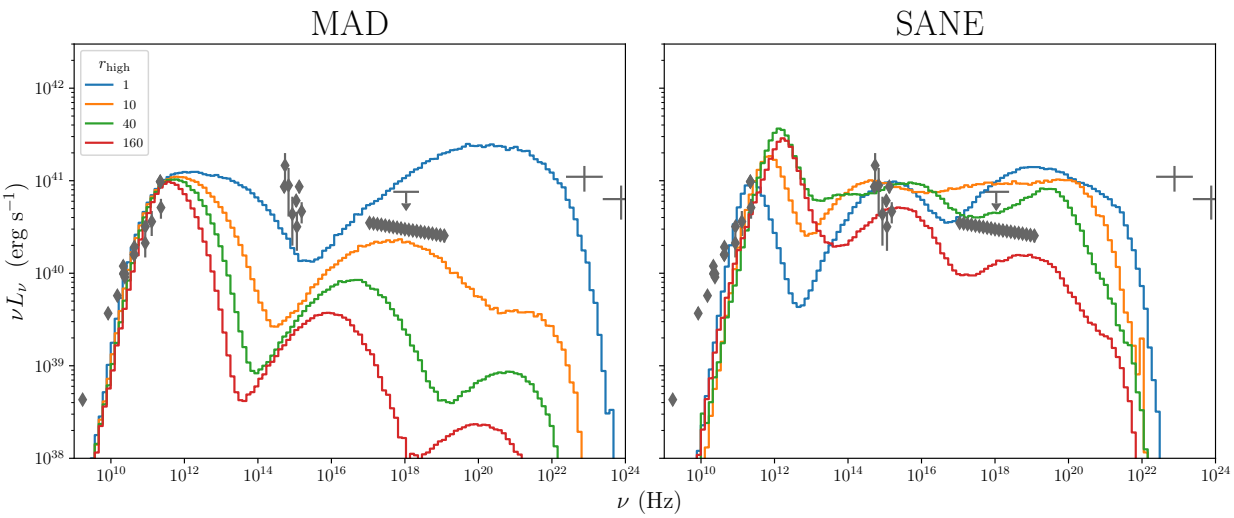


Figure 2.8: Example spectra for MAD (left) and SANE (right) snapshots at low inclination in an $r_{\rm low} = 1$ model. The spectra for $r_{\rm high} = 1, 10, 40, 160$ are included. All data points are taken from the simultaneous multi-wavelength measurement campaign performed coincident with the 2017 EHT observations of M87 and reported in The EHT MWL Science Working Group et al. (2021).

2.5 Future directions

We now briefly discuss future directions as well as improvements and modifications that can be made to the `PATOKA` pipeline.

2.5.1 Radiative transfer model

`ipole` produces images calculated at a single frequency, which neglects the observing bandwidth of the instrument. Extensions to the ray-tracing code could allow for synthesis of finite bandwidth observations; however, this approximation has been found to be inconsequential in the context of the M87 library.

Our treatment generally assumes that the electron distribution function is thermal, i.e., that it is well described by a Maxwell–Jüttner distribution. This assumption enters through definition of the transfer coefficients, which are calculated from the underlying distribution function (see, e.g., Shcherbakov 2008; Pandya et al. 2016, 2018). The introduction of non-thermal electrons can change both the spectral shape and the image morphology (e.g., Özel

et al. 2000; Yuan et al. 2003; Broderick & Loeb 2009; Chael et al. 2017; Mao et al. 2017; Davelaar et al. 2018).

GRMHD simulations produce snapshots of the fluid at different, discrete Kerr–Schild times. The data we present were generated under the *fast light* approximation, where only a single snapshot is used to generate an image or spectrum. This approximation is invalid if the fluid evolves on timescales shorter than the simulation light-crossing time or if one wishes to simulate various time-dependent phenomena, like light echoes from flares or glimmer (e.g., Broderick & Loeb 2005; Moriyama & Mineshige 2015; Wong 2021; Hadar et al. 2021).

The alternative *slow light* method relies on a high fluid snapshot cadence, has large data storage requirements, and has a high throughput cost, which is associated with ray tracing through different time slices. Although not presented here, several slow light simulations were generated to confirm that the fast light approximation does not seriously affect the library results.

2.5.2 Radiative effects

The pipeline we have described used ideal GRMHD to generate the fluid simulations; we have assumed that M87 can be described by models in which radiative cooling is negligible so that it does not affect the dynamics of the plasma. This assumption was probed in Event Horizon Telescope Collaboration et al. (2019e), but it is likely that a full radiative treatment of the fluid simulation will be required in future analyses.

The M87 jet funnel may be populated by electron–positron pairs, produced either via cascades (e.g., Beskin et al. 1992; Levinson & Rieger 2011; Broderick & Tchekhovskoy 2015; Hirotani & Pu 2016) or drizzle (Mościbrodzka et al., 2011; Laurent & Titarchuk, 2018; Kimura & Toma, 2020; Wong et al., 2021b). Ideal GRMHD cannot produce unscreened electric fields and therefore cannot track pair cascades. Computing the cross sections for pair drizzle often requires a high resolution sample of the radiation field, so it is expensive to track *in situ* in GRMHD simulations and is often evaluated in a post-processing step.

Future study is warranted to investigate the signatures of pair plasma emission and whether or not pairs can populate the jet.

2.5.3 The accretion model

The library presented in Section 2.4 is not comprehensive. A more dense sampling of spin may enable a better understanding of how spin affects observables, particularly as spin approaches its maximal value $a_* = \pm 1$. The transitory regime in which the magnetic flux increases from the comfortably SANE state toward the MAD state has been explored, but no dense parameter survey yet exists. The tilted-disk scenario merits further attention and study, even though it increases the size of the parameter space.

A detailed study of the convergence properties of GRMHD simulations both with respect to spatial resolution and simulation duration would be valuable, as would a systematic survey of how the initial conditions affect the statistical properties of the fluid evolution and electromagnetic observables.

2.6 Interpolation and the σ cutoff in detail

GRMHD codes are unable to robustly model fluid evolution in regions with high magnetization, so they often rely on limiting ("flooring") procedures to ensure the numerical stability of the algorithm. The flooring procedures introduce extra, artificial mass and energy, so they must be accounted for when performing ray tracing to generate electromagnetic observables. The standard procedure involves masking each fluid snapshot according to a threshold magnetization value $\sigma_{\rm cut}$. One would hope that the choice of $\sigma_{\rm cut}$ does not materially affect the simulated observables, but it has been observed to alter the morphology of images, especially in simulations of hot MAD flows (see, e.g., Chael et al. 2018a).

Unfortunately, depending on the interpolation algorithm used to reconstruct fluid variables at non-zone-centered locations, the use of a σ mask may also introduce a ridge feature

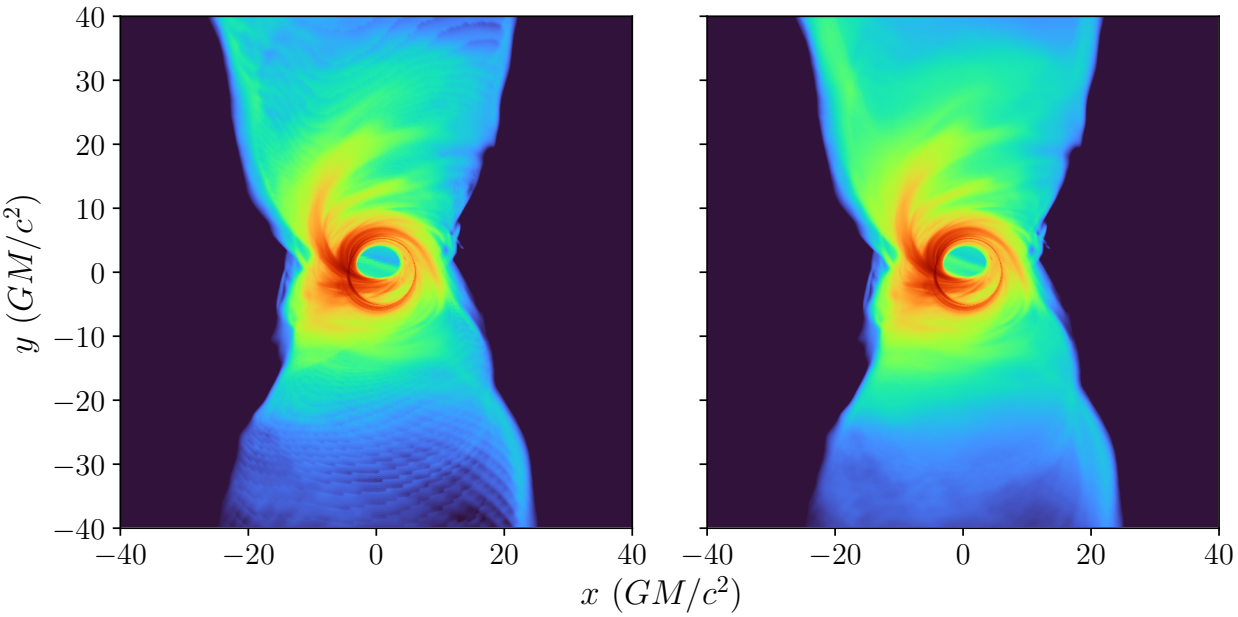


Figure 2.9: Comparison of log-scaled total intensity image using zone-based vs. interpolated-σ cutoffs. The cutoff is implemented by zeroing the electron number density in both panels. The sharp edges are due to interpolation artifacts due to how linear interpolation deals with rapidly varying quantities on a nearly-regular grid. Notice that ridges are still present in the interpolated-σ cutoff image; they are particularly visible near the bottom of the image.

in simulated images that is due to the resolution of the underlying fluid snapshot. Figure 2.9 shows severe (left) and minor (right) examples of the effect. The left panel shows the former default piecewise constant interpolation scheme for σ that zeroed the zone-centered electron number density and introduces sharp boundaries in the image at low intensities. The right panel shows the same image but when σ is (tri-)linearly interpolated and the mask is applied directly to the emissivity at each geodesic step.

Even in the right panel, ridges can still be seen near the bottom of the image. This is due to the linear interpolation scheme, which produces values with discontinuous first derivative at zone centers. In multiple dimensions, the interpolation artifacts are particularly clear: Figure 2.10 shows the interpolated values of a smoothly varying scalar that has been sampled only at zone centers. When the interpolated value is used to mask the transfer coefficients along the geodesics (shown in white in the image), then neighboring geodesics may have noticeably different path lengths.

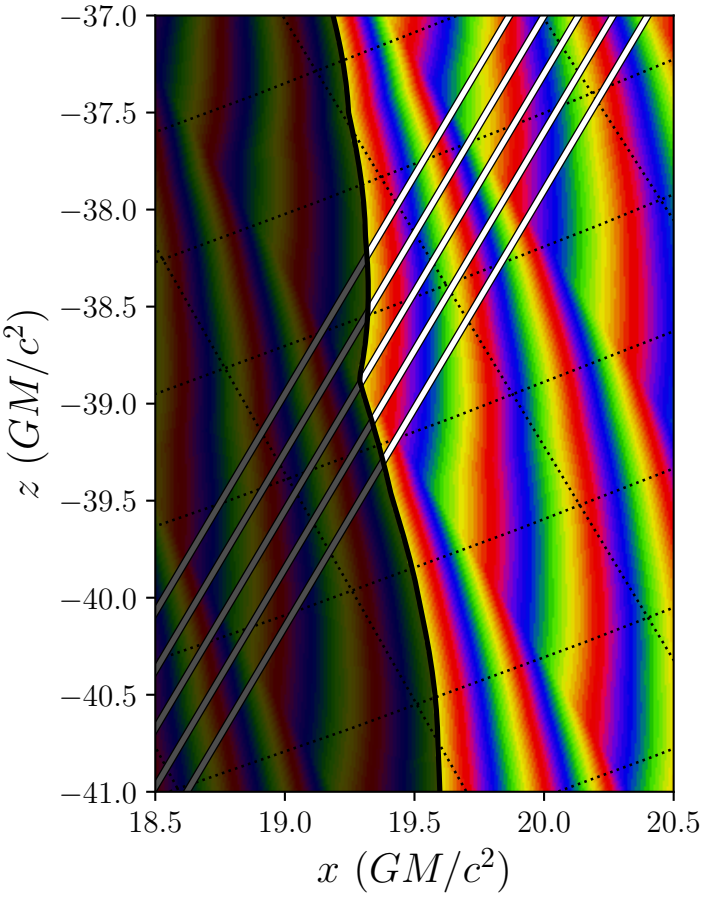


Figure 2.10: Path of five neighboring geodesics (white lines) plotted over bilinearly interpolated magnetization σ (color, rapidly varying rainbow color scale). The $\sigma = 1$ surface is denoted with a solid black line, and zone boundaries are denoted by the dotted black lines. The magnetization decreases steadily from left to right but is defined at zone centers. The value of σ at all other points is reconstructed using bilinear interpolation. The interpolation scheme produces a jagged transition at zone centers. The slight deviations in path length produce ridges even when the value of σ is determined from an interpolated value.

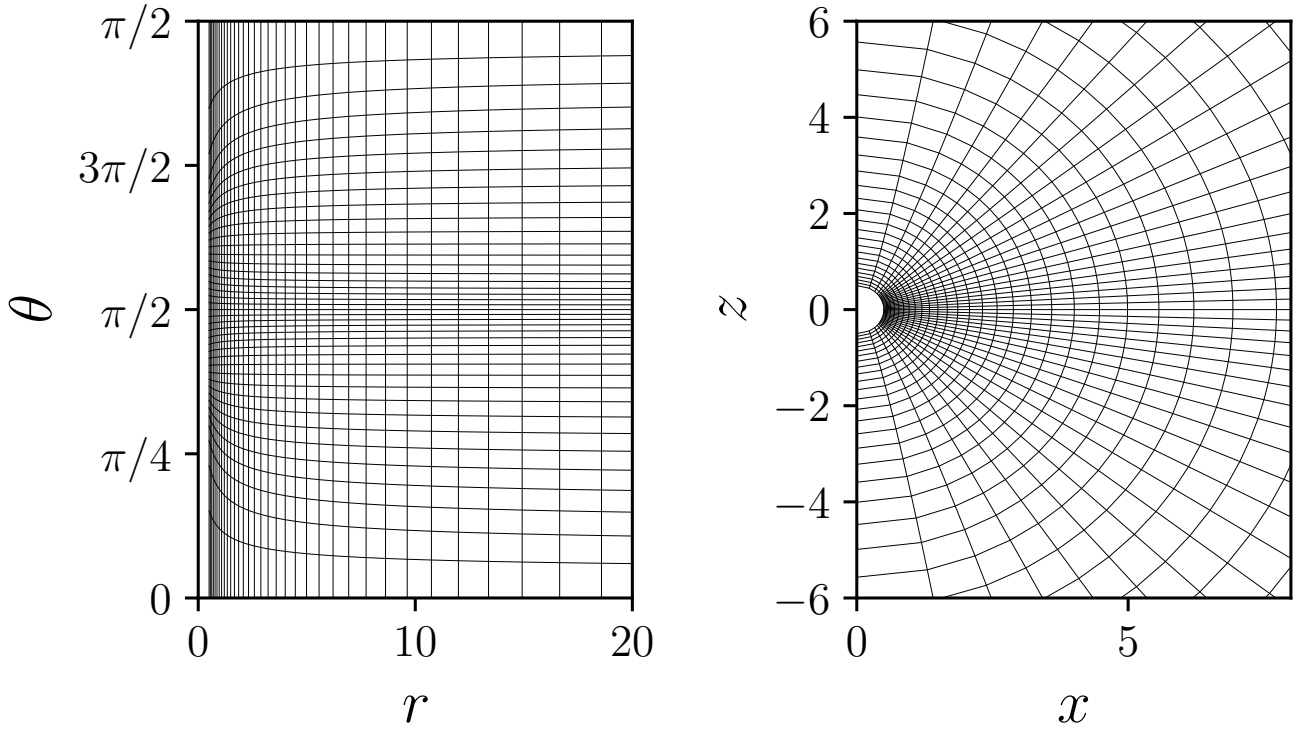


Figure 2.11: Mapping between Kerr–Schild coordinates and funky modified Kerr–Schild coordinates. Left: lines of constant FMKS radial coordinate x^1 (vertical) and latitudinal coordinate x^2 (left-to-right) plotted versus Kerr–Schild radius r and elevation θ. Right: same as left but plotted in a Cartesian embedding with $x = r\sin\theta$ and $z = r\cos\theta$. FMKS coordinates concentrate resolution near the midplane $\theta = \pi/2$ and away from the poles $\theta = 0, \pi$ at small radii.

2.7 FMKS coordinates in detail

Funky modified Kerr–Schild (FMKS) coordinates $x^\mu = (x^0, x^1, x^2, x^3)$ are an extension to the modified Kerr–Schild (MKS) coordinates introduced in Gammie et al. (2003). Positive integer superscripts in this section should be interpreted as indices, not exponents. MKS coordinates are themselves a modification of the horizon-penetrating Kerr–Schild $x^{\overline{\mu}} = (t, r, \theta, \phi)$. Modifications were chosen to both reduce computational cost and increase effective resolution by concentrating zones in regions of the domain where more interesting physics occurs (like the midplane and near the horizon at small radii) and derefining unnecessary small zones. Each of FMKS, MKS, and KS is axisymmetric in ϕ.

Both MKS and FMKS coordinates use an exponential radial coordinate $x^1 \equiv \log(r)$,

which increases the number of zones at small radii where both the relevant dynamical timescale is shorter and it is more important to recover the detailed dynamics of the flow.

FMKS makes two modifications to the elevation coordinate x^2. The first reproduces MKS and increases the number of zones near the midplane by introducing a sinusoidally varying dependence of $\Delta(x^2)$ on θ, as

$$\theta_g \equiv \pi x^2 + \frac{1}{2}(1-h)\sin\left(2\pi x^2\right), \tag{2.34}$$

where h is the midplane "finification" parameter, which we set to $h = 0.3$.

FMKS also introduces a cylindrification in θ whereby zones that are near the poles but are at small radii have larger elevational extent. This choice is meant to increase the required numerical time step, which is set by the minimum of the signal-crossing time over all zones. The signal-crossing time in zones near the funnel often approaches the speed of light, and thus this fact combined with the structure of spherical geometry (which keeps the number of azimuthal zones constant regardless of θ) results in many small zones with fast signal crossing times. Thus, through cylindrification, we increase the size of the smallest zones and similarly gain an increase in time step. The cylindrification is achieved by defining

$$\theta_j = N\left(2x^2-1\right)\left(1+\left(\frac{2x^2-1}{B\left(1+\alpha\right)^{1/\alpha}}\right)^{\alpha}\right)+\pi/2, \tag{2.35}$$

where α and B are parameters and where

$$N = \frac{\pi}{2}\left(1+\frac{B^{-\alpha}}{1+\alpha}\right)^{-1} \tag{2.36}$$

is a normalization term. Finally, the Kerr–Schild colatitudinal coordinate is

$$\theta = \theta_g + \exp\left[-s\Delta x^1\right](\theta_j - \theta_g) \tag{2.37}$$

where $\Delta x^1 = x^1 - \log[r_{\rm in}]$ measures the FMKS distance from the inner edge of the simulation. In our simulations, we take $s = 0.5, B = 0.82$, and $\alpha = 14$.

We do not believe that the above coordinate definition is analytically invertible for $x^\mu(x^{\overline{\mu}})$, so a nonlinear root-finding step may be required to find the FMKS coordinates for a KS event, e.g., when setting the camera position during ray tracing.

Chapter 3

Interpreting M87: Total intensity

In 2017, the Event Horizon Telescope (EHT) observed the compact radio source at the center of the elliptical galaxy M87 at 1.3 mm with extremely high angular resolution. In this chapter, I discuss my contribution to the theoretical analysis of the asymmetric ring that is visible in the 2017 EHT data. I discuss the construction of a large library of GRMHD simulation models and synthetic images and how the library was compared against observed visibilities. The model consistency shows that the observed asymmetric ring and shadow are consistent with earlier predictions of strong gravitational lensing of synchrotron emission from a hot plasma orbiting near the black hole event horizon in a background Kerr spacetime. The ring radius and ring asymmetry depend on black hole mass and spin orientation, respectively, and both are therefore expected to be stable when observed in future EHT campaigns. If the black hole spin and M87's large scale jet are aligned, then the black hole spin vector is pointed away from Earth. I discuss how models of non-spinning black holes are inconsistent with the observations because they do not produce sufficiently powerful jets. At the same time, in those models that produce a sufficiently powerful jet, the latter is shown to be powered through a Blandford–Znajek-like mechanism.

3.1 Introduction

In 1918, the galaxy Messier 87 (M87) was observed by Curtis and found to have "a curious straight ray ... apparently connected with the nucleus by a thin line of matter" (Curtis,

This chapter has been adapted from work with collaborators in the Event Horizon Telescope collaboration that was published as ApJL, Volume 875, Issue 1, L5., 2019.

1918). Curtis's ray is now known to be a jet, extending from sub-pc to several kpc scales, and can be observed across the electromagnetic spectrum, from the radio through γ-rays. Very Long Baseline Interferometry (VLBI) observations that zoom in on the nucleus, probing progressively smaller angular scales at progressively higher frequencies up to 86 GHz by Global mm-VLBI array (GMVA) (e.g., Hada et al., 2016; Walker et al., 2018; Boccardi et al., 2017; Kim et al., 2018), have revealed that the jet emerges from a central core. Models of the stellar velocity distribution imply a mass for the central core $M \approx 6.2 \times 10^9 M_\odot$ at a distance of 16.9 mpc (Gebhardt et al., 2011); models of arcsecond-scale emission lines from ionized gas imply a mass that is lower by about a factor of two (Walsh et al., 2013).

The conventional model for the central object in M87 is a black hole surrounded by a geometrically thick, optically thin, disk accretion flow (e.g., Ichimaru, 1977; Rees et al., 1982; Narayan & Yi, 1994, 1995b; Reynolds et al., 1996). The radiative power of the accretion flow ultimately derives from the gravitational binding energy of the inflowing plasma. There is no consensus model for jet launching, but the two main scenarios are that the jet is a magnetically dominated flow that is ultimately powered by tapping the rotational energy of the black hole (Blandford & Znajek, 1977) and that the jet is a magnetically collimated wind from the surrounding accretion disk (Blandford & Payne, 1982; Lynden-Bell, 2006).

VLBI observations of M87 at frequencies $\geq$ 230 GHz with the Event Horizon Telescope (EHT) resolve angular scales of tens of μas, comparable to that of the event horizon (Doeleman et al., 2012; Akiyama et al., 2015; Event Horizon Telescope Collaboration et al., 2019a,b,c, hereafter EHTC I, II, and III). They therefore have the power to probe the nature of the central object and to test models for jet launching. In addition, EHT observations can constrain the key physical parameters of the system, including the black hole mass and spin, accretion rate, and magnetic flux trapped by accreting plasma in the black hole.

In this paper we adopt the working hypothesis that the central object is a black hole described by the Kerr metric, with mass M and dimensionless spin a_*, $-1 < a_* < 1$. Here $a_* \equiv Jc/GM^2$, where J, G, and c are, respectively, the black hole angular momentum,

gravitational constant, and speed of light. In our convention, $a_* < 0$ implies that the angular momentum of the accretion flow and that of the black hole are anti-aligned. Using general relativistic magnetohydrodynamic (GRMHD) models for the accretion flow and synthetic images of these simulations produced by general relativistic radiative transfer calculations, we test whether the results of the 2017 EHT observing campaign (hereafter EHT2017) are consistent with the black hole hypothesis.

This paper is organized as follows. In Section 3.2 we review salient features of the observations and provide order-of-magnitude estimates for the physical conditions in the source. In Section 3.3 we describe the numerical models. In Section 3.4 we outline our procedure for comparing the models to the data in a way that accounts for model variability. In Section 3.5 we show that many of the models cannot be rejected based on EHT data alone. In Section 3.6 we combine EHT data with other constraints on the radiative efficiency, X-ray luminosity, and jet power and show that the latter constraint eliminates all $a_* = 0$ models. Section 3.7 provides a detailed report of the simulation results and implementation choices. In Section 3.8 we discuss limitations of our models. In Section 3.9 we summarize our results and discuss how further analysis of existing EHT data, future EHT data, and multiwavelength companion observations will sharpen constraints on the models.

3.2 Estimates

Event Horizon Telescope Collaboration et al. (2019d) (hereafter EHTC V) presented images generated from EHT2017 data (for details on the array, 2017 observing campaign, correlation, and calibration see EHTC II; EHTC III). A representative image is reproduced in the left panel of Figure 3.1.

Four features of the image in the left panel of Figure 3.1 play an important role in our analysis: (1) the ringlike geometry, (2) the peak brightness temperature, (3) the total flux density, and (4) the asymmetry of the ring. We now consider each in turn.

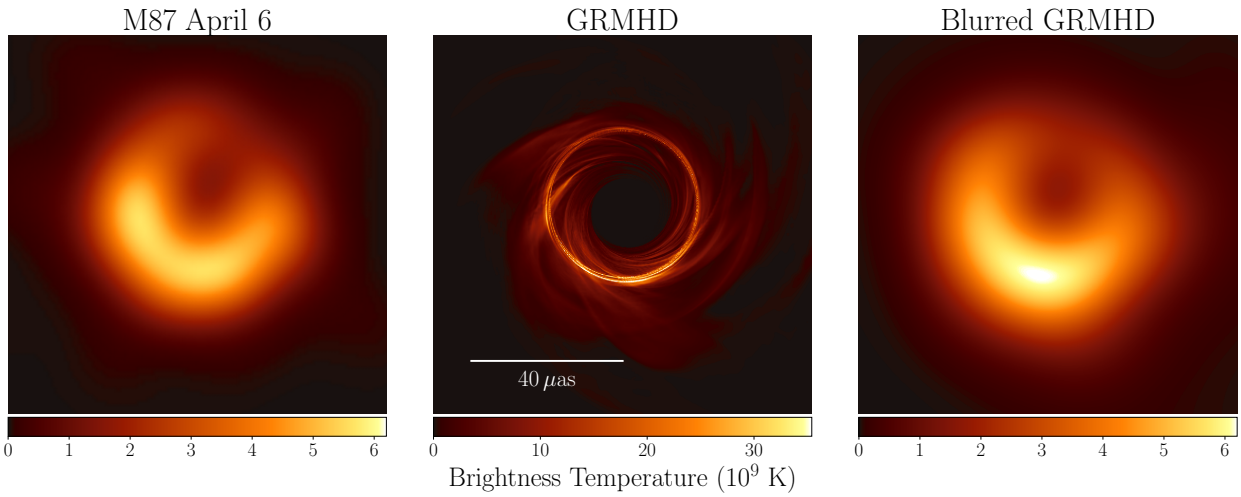


Figure 3.1: Left: An EHT2017 image of M87 from EHTC IV (see their Figure 15). Center: a simulated image based on a GRMHD model. Right: the model image convolved with a 20μas FWHM Gaussian beam. Although the most evident features of the model and data are similar, fine features in the model are not resolved by EHT.

(1) The compact source shows a bright ring with a central dark area without significant extended components. This bears a remarkable similarity to the long-predicted structure for optically thin emission from a hot plasma surrounding a black hole (Falcke et al., 2000). The central hole surrounded by a bright ring arises because of strong gravitational lensing (e.g., Hilbert, 1917; Bardeen et al., 1972; Luminet, 1979). The so-called "photon ring" corresponds to lines of sight that pass close to (unstable) photon orbits (see Teo, 2003), linger near the photon orbit, and therefore have a long path length through the emitting plasma. These lines of sight will appear comparatively bright if the emitting plasma is optically thin. The central flux depression is the so-called black hole "shadow" (Falcke et al., 2000), and corresponds to lines of sight that terminate on the event horizon. The shadow could be seen in contrast to surrounding emission from the accretion flow or lensed counter-jet in M87 (Broderick & Loeb, 2009).

The photon ring is nearly circular for all black hole spins and all inclinations of the black hole spin axis to the line of sight (e.g., Johannsen & Psaltis, 2010). For a $a_* = 0$ black hole

of mass M and distance D, the photon ring angular radius on the sky is

$$\begin{aligned}\theta_p &\equiv \frac{\sqrt{27}\,GM}{c^2 D} \\ &= 18.8 \left(\frac{M}{6.2\times10^9 M_\odot}\right)\left(\frac{D}{16.9\,\mathrm{mpc}}\right)^{-1} \mu\mathrm{as},\end{aligned} \tag{3.1}$$

where we have scaled to the most likely mass from Gebhardt et al. (2011) and a distance of 16.9 mpc (Blakeslee et al., 2009; Bird et al., 2010; Cantiello et al., 2018, see also Event Horizon Telescope Collaboration et al. (2019f), hereafter EHTC VI). The photon ring angular radius for other inclinations and values of a_* differs by at most 13% from equation (3.1), and most of this variation occurs at $1 - |a_*| \ll 1$ (e.g., Takahashi, 2004; Younsi et al., 2016). Evidently the angular radius of the observed photon ring is approximately $\sim 20\mu$as (Fig. 3.1 and EHTC VI), which is close to the prediction of the black hole model given in equation (3.1).

(2) The observed peak brightness temperature of the ring in Figure 3.1 is $T_{b,pk} \sim 6\times10^9$ K, consistent with past EHT mm-VLBI measurements at frequencies ≥ 230 GHz (Doeleman et al., 2012; Akiyama et al., 2015), and GMVA 3 mm-VLBI measurements of the core region (Kim et al., 2018). Expressed in electron rest-mass (m_e) units, $\Theta_{b,pk} \equiv k_\mathrm{B} T_{b,pk}/(m_e c^2) \simeq 1$, where k_B is Boltzmann's constant. The true peak brightness temperature of the source is higher if the ring is unresolved by EHT, as is the case for the model image in the center panel of Figure 3.1.

The 1.3 mm emission from M87 shown in Figure 3.1 is expected to be generated by the synchrotron process (see Yuan & Narayan, 2014, and references therein) and thus depends on the electron distribution function (hereafter eDF). If the emitting plasma has a thermal eDF then it is characterized by an electron temperature $T_e \geq T_b$, or $\Theta_e \equiv k_\mathrm{B} T_e/(m_e c^2) > 1$, since $\Theta_e > \Theta_{b,pk}$ if the ring is unresolved or if the ring is optically thin.

Is the observed brightness temperature consistent with what one would expect from phenomenological models of the source? Radiatively inefficient accretion flow models of M87

(Reynolds et al., 1996; Di Matteo et al., 2003) produce mm emission in a geometrically thick donut of plasma around the black hole. The emitting plasma is collisionless: Coulomb scattering is weak at these low densities and high temperatures. Therefore the electron and ion temperatures need not be the same (e.g., Spitzer, 1962). In radiatively inefficient accretion flow models, the ion temperature is slightly less than the ion virial temperature,

$$\begin{aligned} T_i &\sim 0.3\, T_{i,vir} = 0.3\, m_p c^2 r_{\rm g}/(3k_{\rm B} r) \\ &= 1.1\times 10^{12} (r_{\rm g}/r)\,{\rm K}, \end{aligned} \tag{3.2}$$

where $r_{\rm g} \equiv GM/c^2$ is gravitational radius, r is the Boyer-Lindquist or Kerr–Schild radius, and m_p is the proton mass. Most models have electron temperature $T_e < T_i$ because of electron cooling and preferential heating of the ions by turbulent dissipation (e.g., Yuan & Narayan, 2014; Mościbrodzka et al., 2016). If the emission arises at $\sim 5\, r_{\rm g}$, then $\Theta_e \simeq 37(T_e/T_i)$, which is consistent with the observed $\Theta_{b,pk}$ if the source is unresolved or optically thin.

(3) The total flux density in the image at 1.3 mm is $\simeq 0.5$ Jy. With a few assumptions we can use this to estimate the electron number density n_e and magnetic field strength B in the source. We adopt a simple, spherical, one-zone model for the source with radius $r \simeq 5\, r_{\rm g}$, pressure $n_i k T_i + n_e k T_e = \beta_{\rm p} B^2/(8\pi)$ with $\beta_{\rm p} \equiv p_{\rm gas}/p_{\rm mag} \sim 1$, $T_i \simeq 3T_e$, and temperature $\theta_e \simeq 10\theta_{b,pk}$ consistent with the discussion in (2), above. Setting $n_e = n_i$ (i.e., assuming a fully ionized hydrogen plasma), the values of B and n_e required to produce the observed flux density can be found by solving a nonlinear equation (assuming an average angle between the field and line of sight, 60°). The solution can be approximated as a power-law:

$$\begin{aligned} n_e = 2.9\times 10^4 \left(\frac{r}{5\, r_{\rm g}}\right)^{-1.3} \beta_{\rm p}^{0.62} \times \\ \left(\frac{T_i}{3T_e}\right)^{-0.47} \left(\frac{\theta_e}{10\theta_{b,pk}}\right)^{-2.4} {\rm cm}^{-3} \end{aligned} \tag{3.3}$$

$$B = 4.9 \left(\frac{r}{5\,r_{\rm g}}\right)^{-0.63} \beta_{\rm p}^{-0.19} \times \left(\frac{T_i}{3T_e}\right)^{0.14} \left(\frac{\theta_e}{10\theta_{b,pk}}\right)^{-0.71} \,{\rm G} \tag{3.4}$$

assuming $M = 6.2\times10^9 M_\odot$ and $D = 16.9\,$mpc, and using the approximate thermal emissivity of Leung et al. (2011). Then the synchrotron optical depth at 1.3 mm is ~ 0.2. One can now estimate an accretion rate from (3.3) using

$$\begin{aligned} \dot{M} &= 4\pi r^2 \rho v^r \\ &\sim 4\pi (5r_{\rm g})^2\, n_e m_p\, (c/\sqrt{5}) \\ &\sim 2.7\times10^{-3} M_\odot\,{\rm yr}^{-1} \end{aligned} \tag{3.5}$$

assuming spherical symmetry. The Eddington accretion rate is

$$\dot{M}_{\rm Edd} = \frac{L_{\rm Edd}}{\epsilon c^2} = \frac{2.2}{\epsilon}\left(\frac{M}{10^9\,M_\odot}\right)\, M_\odot\,{\rm yr}^{-1}, \tag{3.6}$$

where $L_{\rm Edd} \equiv 4\pi GMcm_p/\sigma_T$ is the Eddington luminosity (σ_T is Thomson cross section). Setting the efficiency $\epsilon = 0.1$ and $M = 6.2\times10^9 M_\odot$, we find that $\dot{M}_{\rm Edd} = 137 M_\odot\,{\rm yr}^{-1}$, so $\dot{M}/\dot{M}_{\rm Edd} \sim 2.0\times10^{-5}$. This estimate is similar to but slightly larger than the upper limit inferred from the 230 GHz linear polarization properties of M87 (Kuo et al., 2014).

(4) The ring is brighter in the South than the North. This can be explained by a combination of motion in the source and Doppler beaming. As a simple example we consider a luminous, optically thin ring rotating with speed v and with angular momentum vector inclined at a viewing angle $i > 0^\circ$ to the line of sight. Then the approaching side of the ring is Doppler boosted, and the receding side is Doppler dimmed, producing a surface brightness contrast of order unity if v is relativistic. The approaching side of the large scale jet in M87 is oriented West-NorthWest (position angle PA $\approx 288^\circ$; in EHTC VI this is called $\rm PA_{FJ}$), or to the right and slightly up in the image. Walker et al. (2018) estimate that the angle between

the approaching jet and the line of sight is 17°. If the emission is produced by a rotating ring with angular momentum vector oriented along the jet axis, then the plasma in the South is approaching us and the plasma in the North is receding. This implies a clockwise circulation of the plasma in the source, as projected onto the plane of the sky. This sense of rotation is consistent with the sense of rotation in ionized gas at arcsecond scales (Harms et al., 1994; Walsh et al., 2013). Notice that the asymmetry of the ring is consistent with the asymmetry inferred from 43 GHz observations of the brightness ratio between the North and South sides of the jet and counter-jet (Walker et al., 2018).

All of these estimates present a picture of the source that is remarkably consistent with the expectations of the black hole model and with existing GRMHD models (e.g., Dexter et al., 2012; Mościbrodzka et al., 2016). They even suggest a sense of rotation of gas close to the black hole. A quantitative comparison with GRMHD models can reveal more.

3.3 The models

Consistent with the discussion in Section 3.2, we now adopt the working hypothesis that M87 contains a turbulent, magnetized accretion flow surrounding a Kerr black hole. To test this hypothesis quantitatively against the EHT2017 data we have generated a Simulation Library of three-dimensional time-dependent ideal GRMHD models. To generate this computationally expensive library efficiently and with independent checks on the results we used several different codes that evolved matching initial conditions using the equations of ideal GRMHD. The codes used include `BHAC` (Porth et al., 2017), `H-AMR` (Liska et al., 2018; Chatterjee et al., 2019), `iharm` (Gammie et al., 2003), and `KORAL` (Sądowski et al., 2013b, 2014). A comparison of these and other GRMHD codes can be found in Porth et al. (2019), which shows that the differences between integrations of a standard accretion model with different codes is smaller than the fluctuations in individual simulations.

From the Simulation Library we have generated a large Image Library of synthetic images.

Snapshots of the GRMHD evolutions were produced using the general relativistic ray-tracing (GRRT) schemes `ipole` (Mościbrodzka & Gammie, 2018), `RAPTOR` (Bronzwaer et al., 2018), and `BHOSS` (Younsi et al., 2020). A comparison of these and other GRRT codes can be found in Gold et al. (2020).

In the GRMHD models the bulk of the 1.3 mm emission is produced within $\lesssim 10\, r_{\rm g}$ of the black hole, where the models can reach a statistically steady state. It is therefore possible to compute predictive radiative models for this compact component of the source without accurately representing the accretion flow at all radii.

We note that the current state-of-the-art models for M87 are radiation-GRMHD models that include radiative feedback and electron-ion thermodynamics (Ryan et al., 2018; Chael et al., 2019). These models are too computationally expensive for a wide survey of parameter space, so in this paper we only consider nonradiative GRMHD models with a parameterized treatment of the electron thermodynamics.

3.3.1 Simulation library

All GRMHD simulations are initialized with a weakly magnetized torus of plasma orbiting in the equatorial plane of the black hole (e.g., Gammie et al., 2003; De Villiers et al., 2003; McKinney & Blandford, 2009; Porth et al., 2017). We do not consider tilted models, in which the accretion flow angular momentum is misaligned with the black hole spin. The limitations of this approach are discussed in Section 3.8.

The initial torus is driven to a turbulent state by instabilities, including the magnetorotational instability (MRI, see e.g., Balbus & Hawley, 1991). In all cases, the outcome contains a moderately magnetized midplane with orbital frequency comparable to the Keplerian orbital frequency, a corona with gas to magnetic pressure ratio $\beta_{\rm p} \equiv p_{\rm gas}/p_{\rm mag} \sim 1$, and a strongly magnetized region over both poles of the black hole with $B^2/\rho c^2 \gg 1$. We refer to the strongly magnetized region as the funnel, and the boundary between the funnel and the corona as the funnel wall (De Villiers et al., 2005; Hawley & Krolik, 2006). All models in

the library are evolved from $t = 0$ to $t = 10^4\, r_{\rm g} c^{-1}$.

The simulation outcome depends on the initial magnetic field strength and geometry insofar as these affect the magnetic flux through the disk, as discussed below. Once the simulation is initiated, the disk transitions to a turbulent state and loses memory of most other details of the initial conditions. This relaxed turbulent state is found inside a characteristic radius that grows over the course of the simulation. To be confident that we are imaging only those regions that have relaxed, we draw snapshots for comparison with the data from $5 \times 10^3 < t/r_{\rm g} c^{-1} < 10^4$.

Untilted GRMHD models have two key physical parameters. The first is the magnitude of the black hole spin a_*, $-1 < a_* < 1$. The second parameter is the absolute magnetic flux $\Phi_{\rm BH}$ crossing one hemisphere of the event horizon (see Tchekhovskoy et al., 2011; Porth et al., 2019, for a definition). It is convenient to recast $\Phi_{\rm BH}$ in dimensionless form $\phi \equiv \Phi_{\rm BH} \left(\dot{M} {r_{\rm g}}^2 c\right)^{-1/2}$.

The magnetic flux ϕ is nonzero because magnetic field is advected into the event horizon by the accretion flow and sustained by currents in the surrounding plasma. At $\phi > \phi_{max} \sim 15$, [1] numerical simulations show that the accumulated magnetic flux erupts, pushes aside the accretion flow, and escapes (Tchekhovskoy et al., 2011; McKinney et al., 2012). Models with $\phi \sim 1$ are conventionally referred to as SANE (standard and normal evolution; Narayan et al., 2012) models; models with $\phi \sim \phi_{max}$ are conventionally referred to as MAD (magnetically arrested disk; Igumenshchev et al., 2003; Narayan et al., 2003) models.

The Simulation Library contains SANE models with $a_* = -0.94$, -0.5, 0, 0.5, 0.75, 0.88, 0.94, 0.97, and 0.98, and MAD models with $a_* = -0.94$, -0.5, 0, 0.5, 0.75, and 0.94. The Simulation Library occupies 23 TB of disk space and contains a total of 43 GRMHD simulations, with some repeated at multiple resolutions across multiple codes to check consistency (see Porth et al., 2019).

[1] In Heaviside units, where a factor of $\sqrt{4\pi}$ is absorbed into the definition of B, $\phi_{max} \simeq 15$. In the Gaussian units used in some earlier papers, $\phi_{max} \simeq 50$.

3.3.2 Image library generation

To produce model images from the simulations for comparison with EHT observations, we use GRRT to generate a large number of synthetic images and derived VLBI data products. To make the synthetic images we need to specify: (1) the magnetic field, velocity field, and density as a function of position and time; (2) the emission and absorption coefficients as a function of position and time; (3) the inclination angle between the accretion flow angular momentum vector and the line of sight i, the position angle PA, the black hole mass M, and the distance D to the observer. In the following we discuss each input in turn. The reader who is only interested in a high-level description of the Image Library may skip ahead to Section 3.3.3.

(1) GRMHD models provide the absolute velocity field of the plasma flow. Nonradiative GRMHD evolutions are invariant, however, under a rescaling of the density by a factor $\mathscr{M}$. In particular, they are invariant under $\rho \to \mathscr{M}\rho$, field strength $B \to \mathscr{M}^{1/2}B$, and internal energy $u \to \mathscr{M}u$ (the Alfvén speed $B/\rho^{1/2}$ and sound speed $\propto \sqrt{u/\rho}$ are invariant). That is, there is no intrinsic mass scale in a nonradiative model as long as the mass of the accretion flow is negligible in comparison to M.[2] We use this freedom to adjust $\mathscr{M}$ so that the average image from a GRMHD model has 1.3 mm flux density ≈ 0.5 Jy (see EHTC IV). Once $\mathscr{M}$ is set, the density, internal energy, and magnetic field are fully specified.

The mass unit $\mathscr{M}$ determines $\dot{M}$. In our ensemble of models $\dot{M}$ ranges from $2\times10^{-7}\dot{M}_{\rm Edd}$ to $4\times10^{-4}\dot{M}_{\rm Edd}$. Accretion rates vary by model category. The mean accretion rate for MAD models is $\sim 10^{-6}\dot{M}_{\rm Edd}$. For SANE models with $a_* > 0$ it is $\sim 5\times10^{-5}\dot{M}_{\rm Edd}$; and for $a_* < 0$ it is $\sim 2\times10^{-4}\dot{M}_{\rm Edd}$.

(2) The observed radio spectral energy distributions (SEDs) and the polarization characteristics of the source make clear that the 1.3 mm emission is synchrotron radiation, as is typical for active galactic nuclei (AGN). Synchrotron absorption and emission coefficients

[2]For a black hole accreting at the Eddington rate, the ratio of the accreting mass onto a black hole to black hole mass is $\sim 10^{-22}(M/M_\odot)$; in our models mass accretion rate is far below the Eddington rate.

depend on the eDF. In what follows, we adopt a relativistic, thermal model for the eDF (a Maxwell–Jüttner distribution; Jüttner, 1911; Rezzolla & Zanotti, 2013). We discuss the limitations of this approach in Section 3.8.

All of our models of M87 are in a sufficiently low density, high temperature regime that the plasma is collisionless (see Ryan et al. 2018 for a discussion of Coulomb coupling in M87). T_e therefore likely does not equal the temperature provided by the simulations. We set T_e using the GRMHD density ρ, internal energy density u, and plasma $\beta_{\rm p}$ using a simple model:

$$T_e = \frac{2 m_p u}{3 k \rho (2 + R)}, \tag{3.7}$$

where we have assumed that the plasma is composed of hydrogen, that the ions are nonrelativistic, and the electrons are relativistic. Here $R \equiv T_i/T_e$ and

$$R = R_{\rm high} \frac{\beta_{\rm p}^2}{1 + \beta_{\rm p}^2} + \frac{1}{1 + \beta_{\rm p}^2}. \tag{3.8}$$

This prescription has one parameter, $R_{\rm high}$, and sets $T_e \simeq T_i$ in low $\beta_{\rm p}$ regions and $T_e \simeq T_i/R_{\rm high}$ in the midplane of the disk. It is adapted from Mościbrodzka et al. (2016) and motivated by models for electron heating in a turbulent, collisionless plasma that preferentially heats the ions for $\beta_{\rm p} \gtrsim 1$ (e.g., Howes, 2010; Kawazura et al., 2019).

(3) We must specify the observer inclination i, the orientation of the observer through the position angle PA, the black hole mass M, and the distance D to the source. Non-EHT constraints on i, PA, and M are considered below; we have generated images at $i = 12, 17, 22, 158, 163$, and 168° and a few at $i = 148^\circ$. The position angle can be changed by rotating the image. All features of the models that we have examined, including $\dot{M}$, are insensitive to small changes in i. The image morphology does depend on whether i is greater than or less than 90°, as we will show below.

The model images are generated with a $160 \times 160 \mu$as field of view and 1μas pixels, which are small compared to the $\sim 20\mu$as nominal resolution of EHT2017. Our analysis is

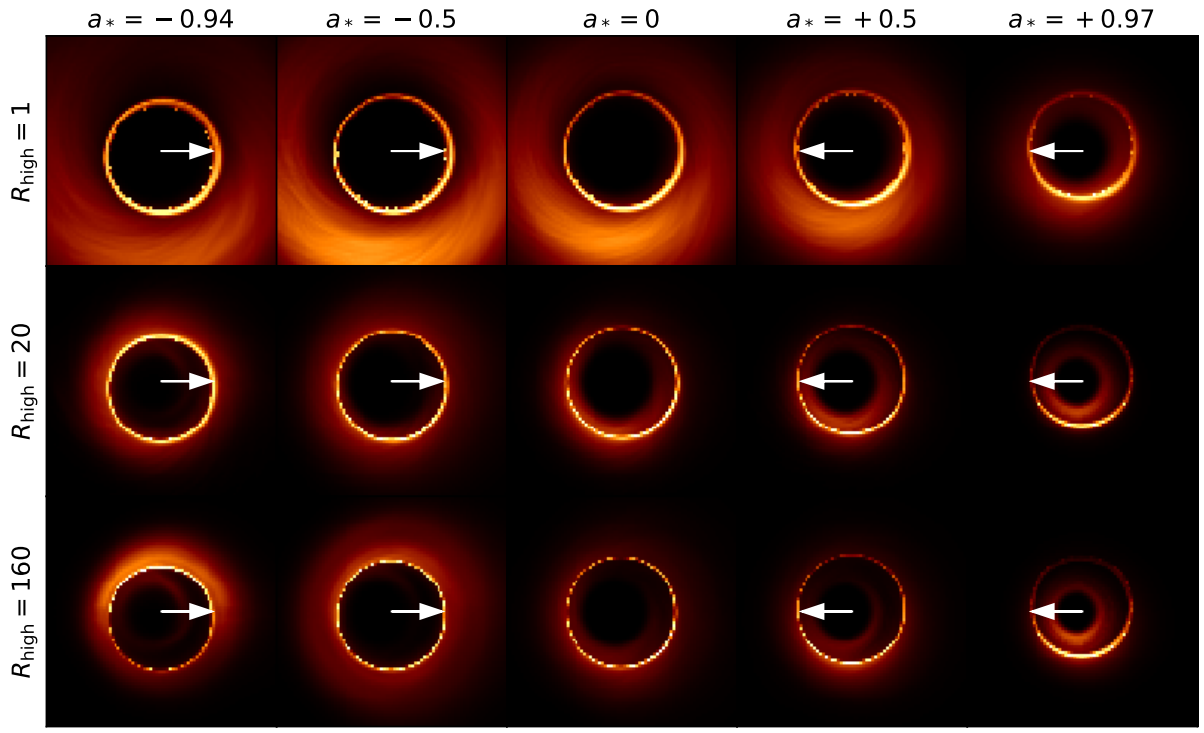


Figure 3.2: Time-averaged 1.3 mm images generated by five SANE GRMHD simulations with varying spin ($a_* = -0.94$ to $a_* = +0.97$ from left to right) and $R_{\rm high}$ ($R_{\rm high} = 1$ to $R_{\rm high} = 160$ from top to bottom; increasing $R_{\rm high}$ corresponds to *decreasing* electron temperature). The colormap is linear. All models are imaged at $i = 163°$. The jet that is approaching us is on the right (West) in all the images. The black hole spin vector projected onto the plane of the sky is marked with an arrow and aligned in the East-West direction. When the arrow is pointing left the black hole rotates in a clockwise direction and when the arrow is pointing right the black hole rotates in a counterclockwise direction. The field-of-view for each model image is 80μas (half of that used for the image libraries) with resolution equal to 1μas/pixel (20 times finer than the nominal resolution of EHT2017, and the same employed in the library images.)

insensitive to changes in the field of view and the pixel scale.

For M we use the most likely value from the stellar absorption-line work, $6.2 \times 10^9 M_\odot$ (Gebhardt et al., 2011). For the distance D we use 16.9 mpc, which is very close to that employed in EHTC VI. The ratio $GM/(c^2 D) = 3.62\mu$as (hereafter M/D) determines the angular scale of the images. We have also generated images with $M = 3.5 \times 10^9 M_\odot$ and confirmed that the results are not predetermined by the input black hole mass.

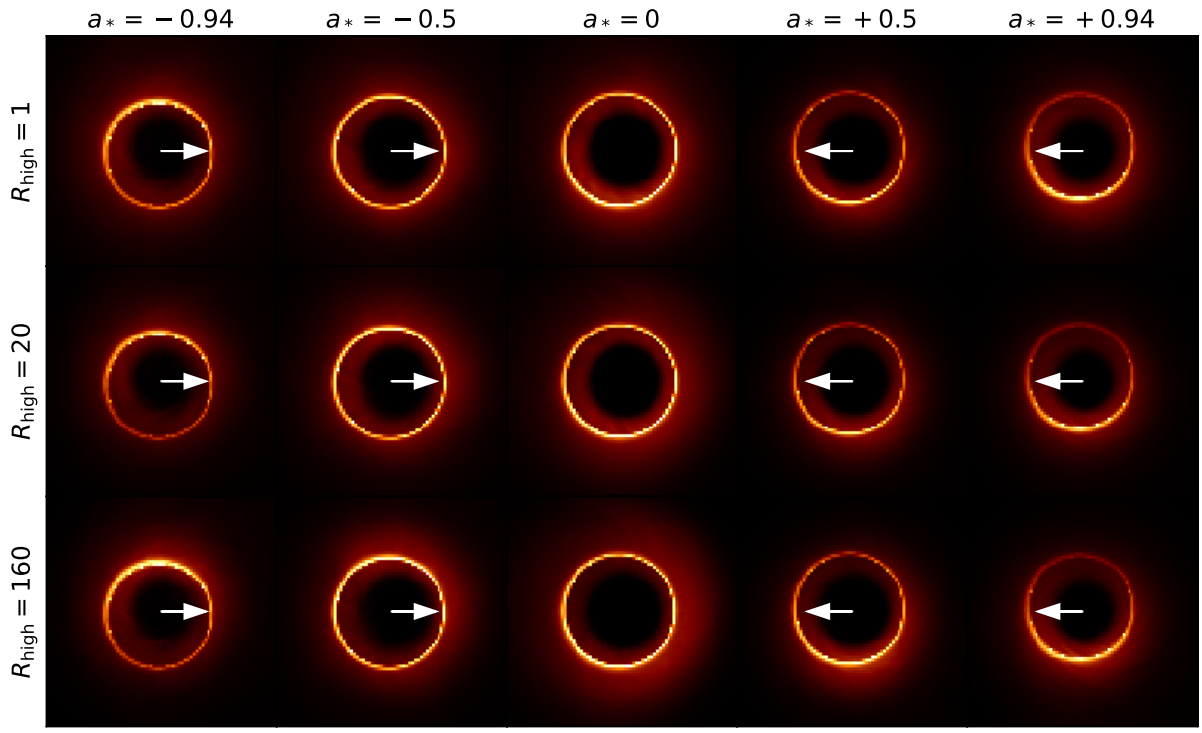


Figure 3.3: Same as in Fig. 3.2 but for selected MAD models.

3.3.3 Image library summary

The Image Library contains of order 60000 images. We generate images from $100 - 500$ distinct output files from each of the GRMHD models at each of $R_{\rm high} = 1, 10, 20, 40, 80$, and 160. In comparing to the data, we adjust the PA by rotation and the total flux and angular scale of the image by rescaling images from the standard parameters in the Image Library (see Fig. 29 in EHTC VI). Tests indicate that comparisons with the data are insensitive to the rescaling procedure unless the angular scaling factor or flux scaling factor is large.[3] We find that the data comparison results are also insensitive to image resolution.[4]

A representative set of time-averaged images from the Image Library are shown in Fig-

[3]In particular the distribution of best-fit M/D, which is defined in Section 3.4, have mean and standard deviation of $M/D = 3.552 \pm 0.605\mu$as when the images are made with an input $M/D = 3.62\mu$as, and $3.564 \pm 0.537\mu$as when the images are made with an input $M/D = 2.01\mu$as. We have also checked images made with an input 1.3 mm flux ranging from 0.1 to 1.5Jy and find relative changes in M/D and PA of less than 1%.

[4]Doubling the image resolution changes the mean best-fit M/D by 7 nano-arcsec and the best-fit PA by $\sim 0.3°$.

ures 3.2 and 3.3. From these figures it is clear that varying the parameters a_*, ϕ, and $R_{\rm high}$ can change the width and asymmetry of the photon ring and introduce additional structures exterior and interior to the photon ring.

The location of the emitting plasma is shown in Figure 3.5, which shows a map of time- and azimuth-averaged emission regions for four representative $a_* > 0$ models. For SANE models, if $R_{\rm high}$ is low (high), emission is concentrated more in the disk (funnel wall), and the bright section of the ring is dominated by the disk (funnel wall)[5].

Figures 3.2 and 3.3 show that for both MAD and SANE models the bright section of the ring, which is generated by Doppler beaming, shifts from the top for negative spin, to a nearly symmetric ring at $a_* = 0$, to the bottom for $a_* > 0$ (except the SANE $R_{\rm high} = 1$ case, where the bright section is always at the bottom when $i > 90°$). That is, the location of the peak flux in the ring is controlled by the black hole spin: it always lies roughly ninety degrees counterclockwise from the projection of the spin vector on the sky. Some of the ring emission originates in the funnel wall at $r \lesssim 8\,r_{\rm g}$. The rotation of plasma in the funnel wall is in the same sense as plasma in the funnel, which is controlled by dragging of magnetic field lines by the black hole. The funnel wall thus rotates opposite the accretion flow if $a_* < 0$. This effect is discussed in Wong et al. (2021a). The resulting relationships between disk angular momentum, black hole angular momentum, and observed ring asymmetry are illustrated in Figure 3.4.

The time-averaged MAD images are almost independent of $R_{\rm high}$ and depend mainly on a_*. In MAD models much of the emission arises in regions with $\beta_{\rm p} \sim 1$ where $R_{\rm high}$ has little influence over the electron temperature, so the insensitivity to $R_{\rm high}$ is natural (see Figure 3.5). In SANE models emission arises at $\beta_{\rm p} \sim 10$, so the time-averaged SANE images, by contrast, depend strongly on $R_{\rm high}$. In low $R_{\rm high}$ SANE models, extended emission outside the photon ring, arising near the equatorial plane, is evident at $R_{\rm high} = 1$. In large $R_{\rm high}$ SANE models the inner ring emission arises from the funnel wall and once again the image

[5]In GRMHD models the jet core is effectively empty and the density is set by numerical "floors." In our radiative transfer calculations, emission from regions with $B^2/\rho > 1$ is explicitly set to zero.

looks like a thin ring (see Figure 3.5).

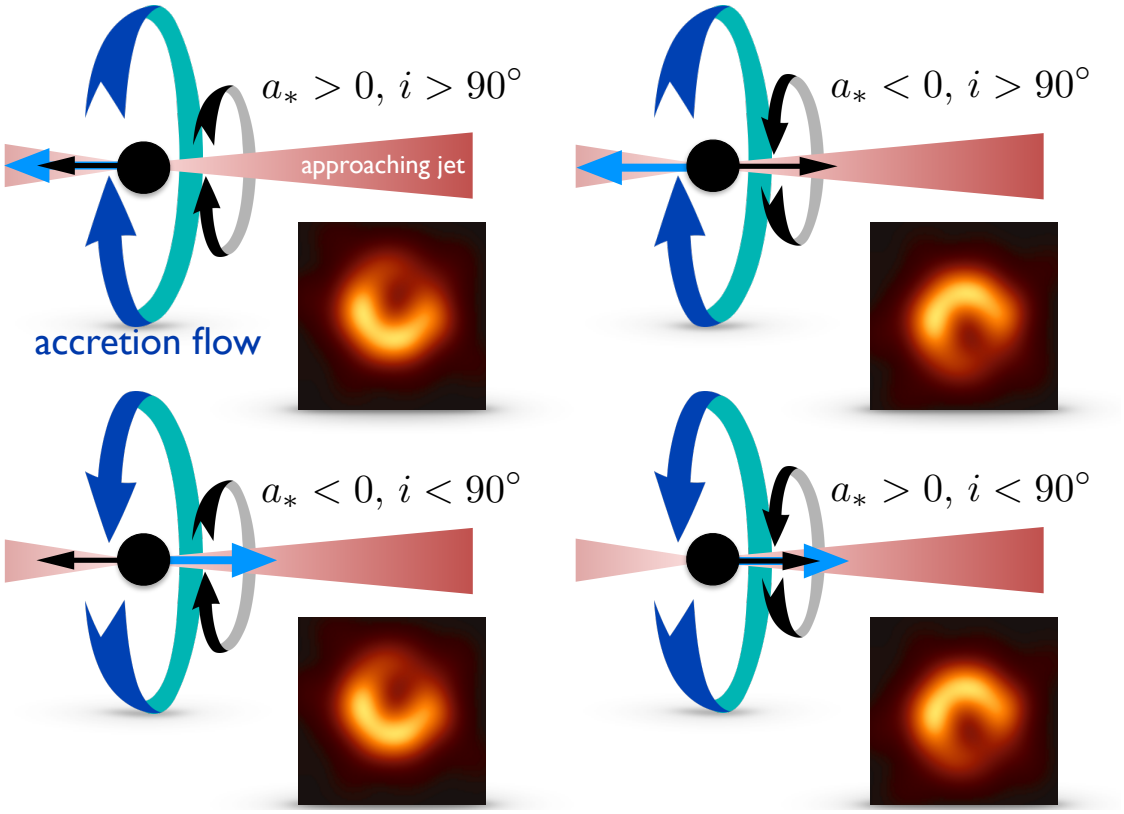


Figure 3.4: Illustration of the effect of black hole and disk angular momentum on ring asymmetry. The asymmetry is produced primarily by Doppler beaming: the bright region corresponds to the approaching side. In GRMHD models that fit the data comparatively well, the asymmetry arises in emission generated in the funnel wall. The sense of rotation of both the jet and funnel wall are controlled by the black hole spin. If the black hole spin axis is aligned with the large scale jet, which points to the right, then the asymmetry implies that the black hole spin is pointing away from Earth (rotation of the black hole is clockwise as viewed from Earth). The blue ribbon arrow shows the sense of disk rotation, and the black ribbon arrow shows black hole spin. Inclination i is defined as the angle between the disk angular momentum vector and the line of sight.

Figure 3.6 shows an example image, set of visibility amplitudes, and closure phases from one of the simulations. An animation of the figure over a $5000\, r_g c^{-1} \approx 5$ yr interval is available in the online version of Event Horizon Telescope Collaboration et al. (2019e). Turbulence in the simulations produces large fluctuations in the images, which produce changes in visibility amplitudes and closure phases that are large compared to measurement errors. The fluctuations are central to our procedure for comparing models with the data,

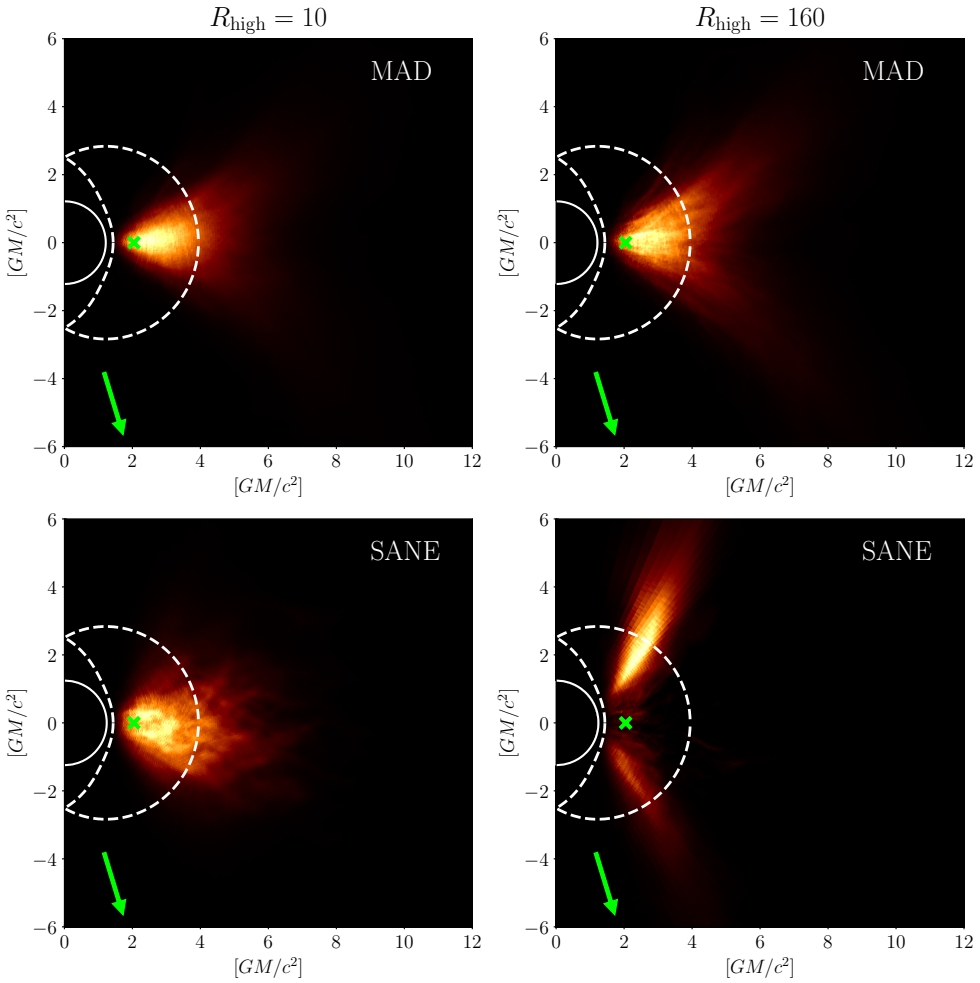


Figure 3.5: Binned location of the point of origin for all photons that make up an image, summed over azimuth, and averaged over all snapshots from the simulation. The colormap is linear. The event horizon is indicated by the solid white semicircle and the black hole spin axis is along the figure vertical axis. This set of four images shows MAD and SANE models with $R_{\rm high} = 10$ and 160, all with $a_* = 0.94$. The region between the dashed curves is the locus of existence of (unstable) photon orbits (Teo, 2003). The green cross marks the location of the ISCO in the equatorial plane. In these images the line of sight (marked by an arrow) is located below the midplane and makes a 163° angle with the disk angular momentum, which coincides with the spin axis of the black hole.

described briefly below and in detail in EHTC VI.

The timescale between different simulation images is $50\, r_{\rm g} c^{-1} \simeq 18$ days, which is long compared to EHT2017 observing campaign. The images are highly correlated on timescales less than the innermost stable circular orbit (ISCO) orbital period, which for $a_* = 0$ is $\simeq 15\ r_{\rm g} c^{-1} \simeq 5$ days, i.e., comparable to the duration of the EHT2017 campaign. If drawn

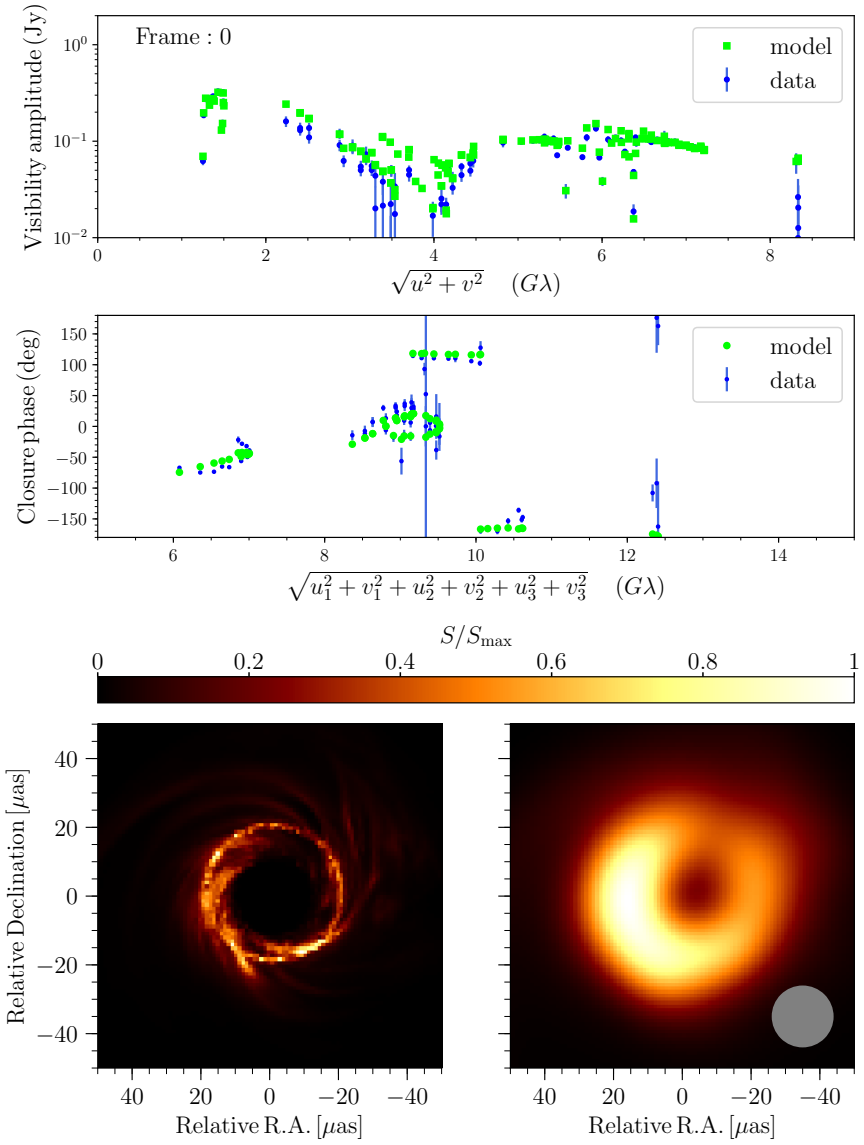


Figure 3.6: Single frame from the accompanying animation. This shows the visibility amplitudes (top), closure phases plotted by Euclidean distance in 6 dimensional space (middle), and associated model images at full resolution (lower left) and convolved with the EHT2017 beam (lower right). Data from April 6th high-band are also shown in the top two plots.

from one of our models, we would expect the EHT2017 data to look like a single snapshot (Figure 3.6) rather than their time averages (Figures 3.2 and 3.3).

3.4 Comparing models with the data

Each model in the Simulation Library has two dimensionless parameters: black hole spin a_* and magnetic flux ϕ. Imaging the model from each simulation adds five new parameters: $R_{\rm high}$, i, PA, M, and D, which we set to 16.9 mpc. After fixing these parameters we draw snapshots from the time evolution at a cadence of 10 to $50\,r_{\rm g}c^{-1}$. We then compare these snapshots to the data.

The simplest comparison computes the χ^2_ν (reduced chi square) distance between the data and a snapshot. In the course of computing χ^2_ν we vary the image scale M/D, flux density F_ν, position angle PA, and the gain at each VLBI station in order to give each image every opportunity to fit the data. The best-fit parameters $(M/D, F_\nu, \mathrm{PA})$ for each snapshot are found by two pipelines independently: the THEMIS pipeline using a Markov chain Monte Carlo (MCMC) method (Broderick et al., 2020), and the `GENA` pipeline using an evolutionary algorithm for multidimensional minimization (see Fromm et al. 2019 and §4 of EHTC VI for details). The best-fit parameters contain information about the source, and we use the distribution of best-fit parameters to test the model by asking whether they are consistent with existing measurements of M/D and estimates of the jet PA on larger scales.

The χ^2_ν comparison alone does not provide a sharp test of the models. Fluctuations in the underlying GRMHD model, combined with the high signal-to-noise ratio for EHT2017 data, imply that individual snapshots are highly unlikely to provide a formally acceptable fit with $\chi^2_\nu \simeq 1$. This is borne out in practice with the minimum $\chi^2_\nu = 1.79$ over the entire set of the more than 60000 individual images in the Image Library. Nevertheless, it is possible to test if the χ^2_ν from the fit to the data is consistent with the underlying model, using "Average Image Scoring" with THEMIS (THEMIS-AIS), as described in detail in Appendix F of EHTC VI. THEMIS-AIS measures a χ^2_ν distance (on the space of visibility amplitudes and closure phases) between a trial image and the data. In practice we use the average of the images from a given model as the trial image (hence THEMIS-AIS), but other choices are possible. We compute the χ^2_ν distance between the trial image and synthetic data produced

from each snapshot. The model can then be tested by asking whether the data's χ^2_ν is likely to have been drawn from the model's distribution of χ^2_ν. In particular, we can assign a probability p that the data is drawn from a specific model's distribution.

In this paper we focus on comparisons with a single dataset, the 2017 April 6 high-band data (EHTC III). The eight EHT2017 datasets, spanning four days with two bands on each day, are highly correlated. The 2017 April 6th dataset has the largest number of scans, 284 detections in 25 scans (see EHTC III) and is therefore expected to be the most constraining.[6]

3.5 Model constraints: EHT2017 alone

The resolved ringlike structure obtained from the EHT2017 data provides an estimate of M/D (discussed in detail in EHTC VI) and the jet PA from the immediate environment of the central black hole. As a first test of the models we can ask whether these are consistent with what is known from other mass measurements and from the orientation of the large scale jet.

Figure 3.7 shows the distributions of best-fit values of M/D for a subset of the models for which spectra and jet power estimates are available (see below). The three lines show the M/D distribution for all snapshots (dotted lines), the best fit 10% of snapshots (dashed lines), and the best fit 1% of snapshots (solid lines) within each model. Evidently, as better fits are required, the distribution narrows and peaks close to $M/D \sim 3.6\mu$as with a width of about 0.5μas.

The distribution of M/D for the best fit $< 10\%$ of snapshots is qualitatively similar if we include only MAD or SANE models, only models produced by individual codes (`BHAC`, `H-AMR`, `iharm`, or `KORAL`), or only individual spins. Since the thrust of this paper is to test the models, we simply note that Figure 3.7 indicates that the models are broadly consistent with earlier mass estimates (see EHTC VI for a detailed discussion), although this did not have to be the case.

[6]EHTC I and EHTC IV focus instead on the April 11 dataset.

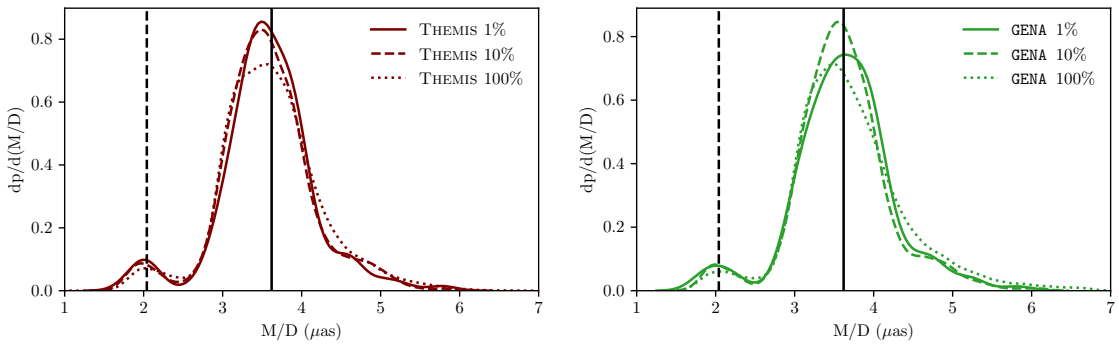


Figure 3.7: Distribution of M/D obtained by fitting Image Library snapshots to the April 6 data, in μas, measured independently using the (left) THEMIS and (right) GENA pipelines with qualitatively similar results. Smooth lines were drawn with a Gaussian kernel density estimator. The three lines show (solid) the best-fit 1% within each model; (dashed) the best-fit 10% within each model; (dotted) all model images. The vertical lines show $M/D = 2.04$ (dashed) and 3.62μas (solid), corresponding to $M = 3.5$ and $6.2 \times 10^9 M_\odot$. The distribution uses a subset of models for which spectra and jet power estimates are available (see Section 3.6). Only images with $a_* > 0$, $i > 90°$ and $a_* < 0$, $i < 90°$ (see also left panel of figure 3.4) are considered.

We can go further and ask if any of the individual models favor large or small masses. Figure 3.8 shows the distributions of best-fit values of M/D for each model (different a_*, $R_{\rm high}$, and magnetic flux). Most individual models favor M/D close to 3.6μas. The exceptions are $a_* \leq 0$ SANE models with $R_{\rm high} = 1$, which produce the bump in the M/D distribution near 2μas. In these models, the emission is produced at comparatively large radius in the disk (see Figure 3.2) because the inner edge of the disk (the ISCO) is at large radius in a counterrotating disk around a black hole with $|a_*| \sim 1$. For these models, the fitting procedure identifies EHT2017's ring with this outer ring, which forces the photon ring, and therefore M/D, to be small. As we will show later, these models can be rejected because they produce weak jets that are inconsistent with existing jet power estimates (see Section 3.6.3).

Figure 3.8 also shows that M/D increases with a_* for SANE models. This is due to the appearance of a secondary inner ring inside the main photon ring. The former is associated with emission produced along the wall of the approaching jet. Since the emission is produced in front of the black hole lensing is weak and it appears at small angular scale. The inner

ring is absent in MAD models (see Figure 3.3), where the bulk of the emission comes from the midplane at all values of $R_{\rm high}$ (Figure 3.5).

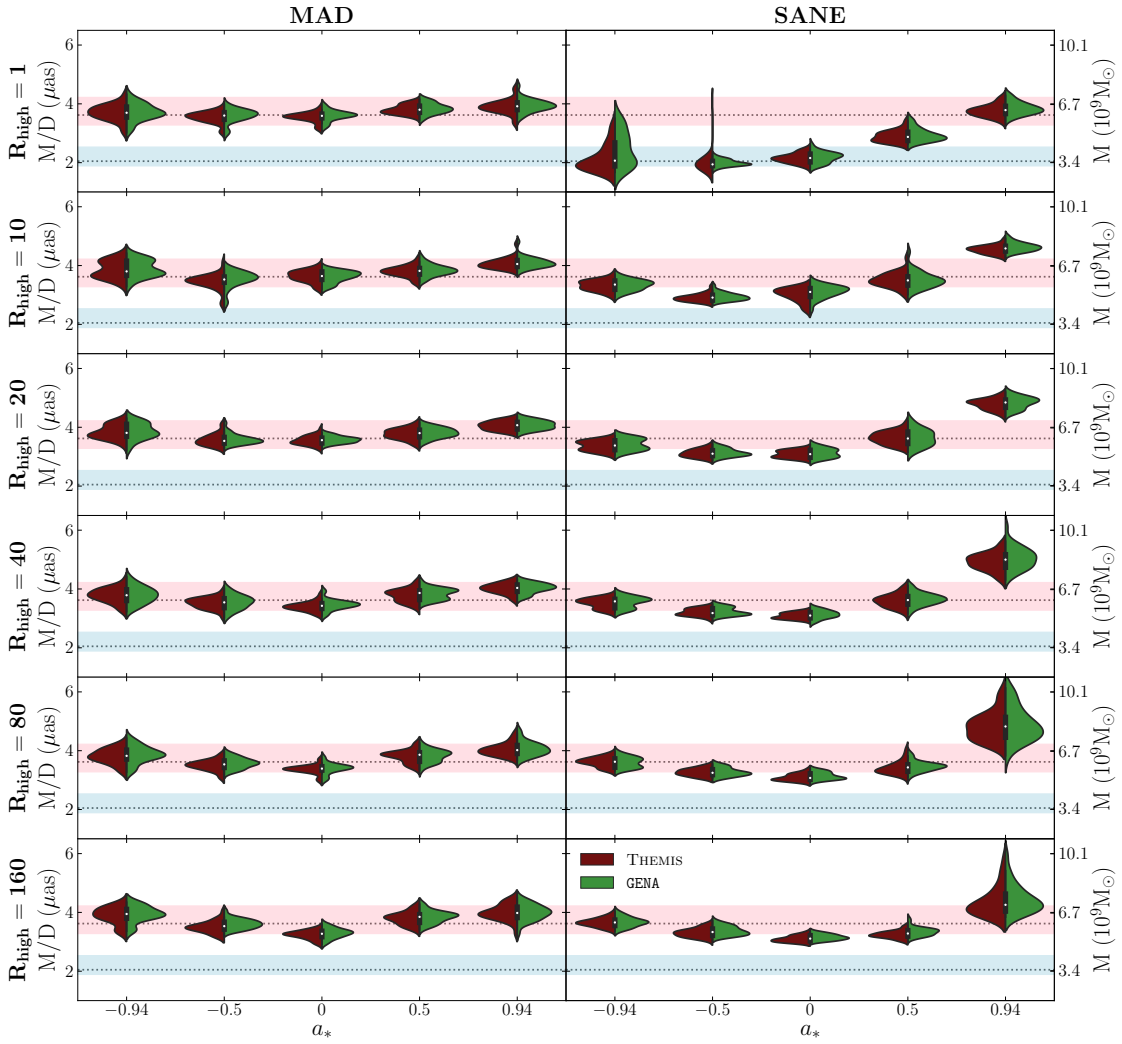


Figure 3.8: Distributions of M/D and black hole mass with $D = 16.9\,$mpc reconstructed from the best-fit 10% of images for MAD (left panel) and SANE (right panel) models ($i = 17°$ for $a_* > 0$ and 163° for $a_* \leq 0$) with different $R_{\rm high}$ and a_*, from the THEMIS (red, left), and `GENA` (green, right) pipelines. The white dot and vertical black bar correspond respectively to the median and region between the 25th and 75th percentiles for both pipelines combined. The blue and pink horizontal bands show the range of M/D and mass at $D = 16.9\,$mpc estimated from the gas dynamical model (Walsh et al., 2013) and stellar dynamical model (Gebhardt et al., 2011), respectively. Constraints on the models based on average image scoring (THEMIS-AIS) are discussed below. Constraints based on radiative efficiency, X-ray luminosity, and jet power are discussed in Section 3.6.

We now ask whether the position angle of the jet is consistent with the orientation of

the jet measured at other wavelengths. On large ($\sim$ mas) scales the extended jet component has a PA of approximately 288° (e.g., Walker et al., 2018). On smaller ($\sim 100\,\mu$as) scales the apparent opening angle of the jet is large (e.g., Kim et al., 2018) and the PA is therefore more difficult to measure. Also notice that the jet PA may be time dependent (e.g., Walker et al., 2018; Hada et al., 2017). In our model images the jet is relatively dim at 1.3 mm, and is not easily seen with a linear colormap. The model jet axis is, nonetheless, well defined: jets emerge perpendicular to the disk.

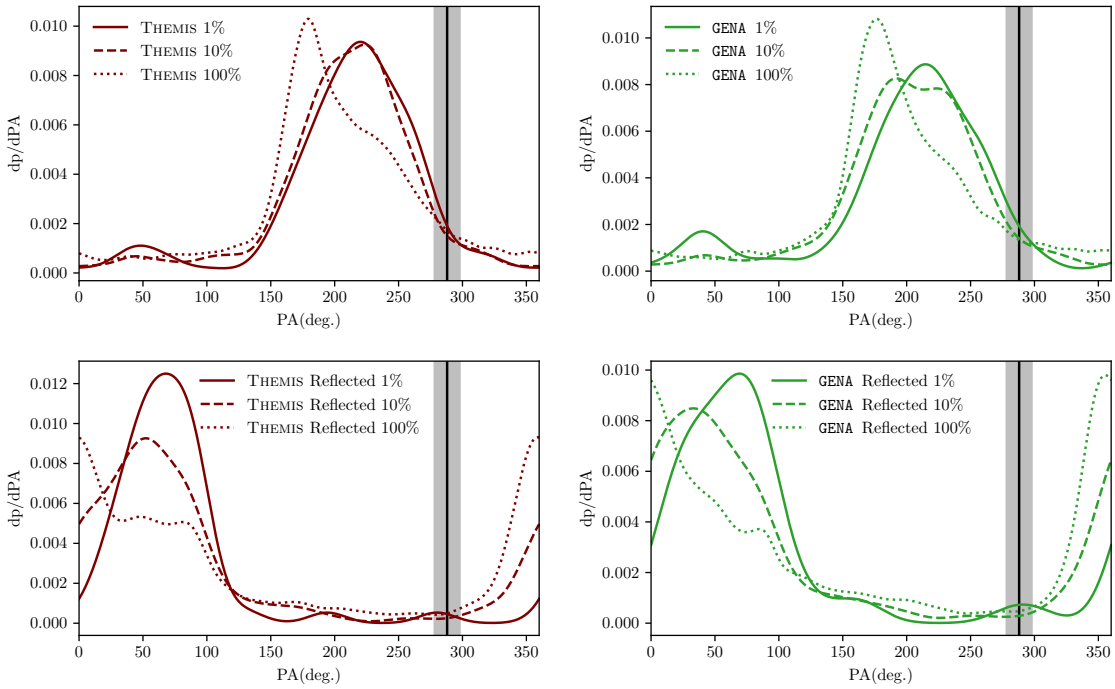


Figure 3.9: Top: Distribution of best fit PA (in degree) scored by the THEMIS (left) and `GENA` (right) pipelines for models with black hole spin vector pointing away from Earth ($i > 90°$ for $a_* > 0$ or $i < 90°$ for $a_* < 0$). Bottom: Images with black hole spin vector pointing towards Earth ($i < 90°$ for $a_* > 0$ or $i > 90°$ for $a_* < 0$). Smooth lines were drawn with a wrapped Gaussian kernel density estimator. The three lines show (1) all images in the sample (dotted line); (2) the best-fit 10% of images within each model (dashed line); and (3) the best-fit 1% of images in each model (solid line). For reference, the vertical line shows the position angle PA $\sim 288°$ of the large scale (mas) jet Walker et al. (2018), with the grey area from $(288 - 10)°$ to $(288 + 10)°$ indicating the observed PA variation.

Figure 3.9 shows the distribution of best fit PA over the same sample of snapshots from the Image Library used in Figure 3.7. We divide the snapshots into two groups. The first

group has the black hole spin pointed away from us ($i > 90°$ and $a_* > 0$, or $i < 90°$ and $a_* < 0$). The spin-away model PA distributions are shown in the top two panels. The second group has the black hole spin pointed toward us ($i > 90$ and $a_* < 0$ or $i > 90°$ and $a_* < 0$). These spin-toward model PA distributions are shown in the bottom two panels. The large scale jet orientation lies on the shoulder of the spin-away distribution (the distribution can be approximated as a Gaussian with, for THEMIS (`GENA`) mean 209 (203)° and $\sigma_{\rm PA} = 54$ (55)°; the large-scale jet PA lies $1.5\sigma_{\rm PA}$ from the mean) and is therefore consistent with the spin-away models. On the other hand, the large scale jet orientation lies off the shoulder of the spin-toward distribution and is inconsistent with the spin-toward models. Evidently models in which the black hole spin is pointing away from us are strongly favored.

The width of the spin-away and spin-toward distributions arises naturally in the models from brightness fluctuations in the ring. The distributions are relatively insensitive if split into MAD and SANE categories, although for MAD the averaged PA is $\langle {\rm PA} \rangle = 219°$, $\sigma_{\rm PA} = 46°$, while for SANE $\langle {\rm PA} \rangle = 195°$ and $\sigma_{\rm PA} = 58°$. The $a_* = 0$ and $a_* > 0$ models have similar distributions. Again, EHT2017 data strongly favor one sense of black hole spin: either $|a_*|$ is small, or the spin vector is pointed away from us. If the fluctuations are such that the fitted PA for each epoch of observations is drawn from a Gaussian with $\sigma_{\rm PA} \simeq 55°$ then a second epoch will be able to identify the true orientation with accuracy $\sigma_{\rm PA}/\sqrt{2} \simeq 40°$ and the N-th epoch with accuracy $\sigma_{\rm PA}/\sqrt{N}$. If the fitted PA were drawn from a Gaussian of width $\sigma_{\rm PA} = 54°$ about PA = 288°, as would be expected in a model in which the large-scale jet is aligned normal to the disk, then future epochs have a $> 90\%$ chance of seeing the peak brightness counterclockwise from its position in EHT2017.

Finally, we can test the models by asking if they are consistent with the data according to THEMIS-AIS, as introduced in Section 3.4. THEMIS-AIS produces a probability p that the χ^2_ν distance between the data and the average of the model images is drawn from the same distribution as the χ^2_ν distance between synthetic data created from the model images and the average of the model images. Table 3.1 takes these p values and categorizes them by

Flux	a_*	$\langle p \rangle$	$N_{\rm model}$	MIN(p)	MAX(p)
SANE	-0.94	0.33	24	0.01	0.88
SANE	-0.5	0.19	24	0.01	0.73
SANE	0	0.23	24	0.01	0.92
SANE	0.5	0.51	30	0.02	0.97
SANE	0.75	0.74	6	0.48	0.98
SANE	0.88	0.65	6	0.26	0.94
SANE	0.94	0.49	24	0.01	0.92
SANE	0.97	0.12	6	0.06	0.40
MAD	-0.94	0.01	18	0.01	0.04
MAD	-0.5	0.75	18	0.34	0.98
MAD	0	0.22	18	0.01	0.62
MAD	0.5	0.17	18	0.02	0.54
MAD	0.75	0.28	18	0.01	0.72
MAD	0.94	0.21	18	0.02	0.50

Table 3.1: Summary results of average image scoring (THEMIS-AIS) procedure introduced in Section 3.4. $\langle p \rangle$ reports the mean p value for the aggregated models, $N_{\rm model}$ is the number of aggregated models, MIN(p) and MAX(p) are the minimum and maximum p values among the set of aggregated models.

magnetic flux and by spin, aggregating (averaging) results from different codes, $R_{\rm high}$, and i. Evidently, most of the models are formally consistent with the data by this test.

One group of models, however, is rejected by THEMIS-AIS: MAD models with $a_* = -0.94$. On average this group has $p = 0.01$, and all models within this group have $p \leq 0.04$. Snapshots from MAD models with $a_* = -0.94$ exhibit the highest morphological variability in our ensemble in the sense that the emission breaks up into transient bright clumps. These models are rejected by THEMIS-AIS because *none* of the snapshots are as similar to the average image as the data is. In other words, it is unlikely that EHT2017 would have captured a $a_* = -0.94$ MAD model in a configuration as unperturbed as the data seem to be.

The remainder of the model categories contain at least some models that are consistent with the data according to the average image scoring test—most models are variable and the associated snapshots lie far from the average image. These snapshots are formally inconsistent with the data, but their distance from the average image is consistent with

what is expected from the models. Given the uncertainties in the model and our lack of knowledge of the source prior to EHT2017, it is remarkable that so many of the models are acceptable. This is likely because the source structure is dominated by the photon ring, which is produced by gravitational lensing and therefore relatively insensitive to the details of the accretion flow and jet physics. We can further narrow the range of acceptable models, however, using additional constraints.

3.6 Model constraints: EHT2017 combined with other constraints

We can apply three additional arguments to further constrain the source model. (1) The model must be close to radiative equilibrium. (2) The model must be consistent with the observed broad-band SED; in particular, it must not overproduce X-rays. (3) The model must produce a sufficiently powerful jet to match the measurements of the jet kinetic energy at large scales. Our discussions in this Section are based on simulation data that is provided in full detail in § 3.7.

3.6.1 Radiative equilibrium

The model must be close to radiative equilibrium. The GRMHD models in the Simulation Library do not include radiative cooling nor do they include a detailed prescription for particle energization. In nature, the accretion flow and jet are expected to be cooled and heated by a combination of synchrotron and Compton cooling, turbulent dissipation, and Coulomb heating, which transfers energy from the hot ions to the cooler electrons. In our suite of simulations the parameter $R_{\rm high}$ can be thought of as a proxy for the sum of these processes. In a fully self-consistent treatment, some models would rapidly cool and settle to a lower electron temperature (see Mościbrodzka et al., 2011; Ryan et al., 2018; Chael et al., 2019). We crudely test for this by calculating the radiative efficiency $\epsilon \equiv L_{\rm bol}/(\dot{M}c^2)$,

where $L_{\rm bol}$ is the bolometric luminosity. If it is larger than the radiative efficiency of a thin, radiatively efficient disk,[7] which depends only on a_* (Novikov & Thorne, 1973), then we reject the model as physically inconsistent.

We calculate $L_{\rm bol}$ with the Monte Carlo code `igrmonty` (Dolence et al., 2009), which incorporates synchrotron emission, absorption, Compton scattering at all orders, and bremsstrahlung. It assumes the same, thermal eDF used in generating the Image Library. We calculate $L_{\rm bol}$ for 20% of the snapshots to minimize computational cost. We then average over snapshots to find $\langle L_{\rm bol} \rangle$. The mass accretion rate $\dot{M}$ is likewise computed for each snapshot and averaged over time. We reject models with ϵ larger than the classical thin disk model—see Table 3.3 in § 3.7 for a list of the ϵ for many of the models. All but two of the radiatively inconsistent models are MADs with $a_* \geq 0$ and $R_{\rm high} = 1$. Eliminating all MAD models with $a_* \geq 0$ and $R_{\rm high} = 1$ does not change any of our earlier conclusions.

3.6.2 X-ray constraints

As part of the EHT2017 campaign, we simultaneously observed M87 with the *Chandra* X-ray observatory and the Nuclear Spectroscopic Telescope Array (*NuSTAR*). The best fit to simultaneous *Chandra* and *NuSTAR* observations on April 12 and 14 implies a $2-10\,{\rm keV}$ luminosity of $L_{\rm Xobs} = 4.4 \pm 0.1 \times 10^{40}\,{\rm erg\,sec^{-1}}$. We used the SEDs generated from the simulations while calculating $L_{\rm bol}$ to reject models that consistently overproduce X-rays; specifically, we reject models with $\log L_{\rm Xobs} < \log\langle L_{\rm X} \rangle - 2\sigma(\log L_{\rm X})$. We do not reject underluminous models since the X-rays could in principle be produced by direct synchrotron emission from non-thermal electrons or by other unresolved sources. Notice that $L_{\rm X}$ is highly variable in all models so the X-ray observations only reject a few models. Table 3.3 in § 3.7 shows $\langle L_{\rm X} \rangle$ as well as upper and lower limits for a set of models that is distributed uniformly across the parameter space.

[7]The thin disk radiative efficiency is 0.038 for $a_* = -1$, 0.057 for $a_* = 0$, and 0.42 for $a_* = 1$. See equations (2.12) and (2.21) of Bardeen et al. (1972); the efficiency is $1 - E/\mu_p$, where μ_p is the rest mass of the particle. The rejected model list is identical if instead one simply rejects all models with $\epsilon > 0.2$.

The X-ray flux in our models is produced by inverse Compton scattering of synchrotron photons. The X-ray flux is an increasing function of $\tau_T T_e^2$ where τ_T is a characteristic Thomson optical depth ($\tau_T \sim 10^{-5}$), and the characteristic amplification factor for photon energies is $\propto T_e^2$ because the X-ray band is dominated by singly scattered photons interacting with relativistic electrons (we include all scattering orders in the Monte Carlo calculation). Increasing $R_{\rm high}$ at fixed $F_\nu(230\,{\rm GHz})$ tends to increase $\dot{M}$ (and therefore τ_T) and decrease T_e. The increase in T_e dominates in our ensemble of models and so models with small $R_{\rm high}$ have larger $L_{\rm X}$, while models with large $R_{\rm high}$ have smaller $L_{\rm X}$. The effect is not strictly monotonic, however, because of noise in our sampling process and the highly variable nature of the X-ray emission.

The overluminous models are mostly SANE models with $R_{\rm high} \leq 20$. The model with highest $\langle L_{\rm X} \rangle = 4.2 \times 10^{42}\,{\rm erg\,sec^{-1}}$ is a SANE, $a_* = 0$, $R_{\rm high} = 10$ model. The corresponding model with $R_{\rm high} = 1$ has $\langle L_{\rm X} \rangle = 2.1 \times 10^{41}\,{\rm erg\,sec^{-1}}$, and the difference between these two indicates the level of variability and the sensitivity of the average to the brightest snapshot. The upshot of application of the $L_{\rm X}$ constraints is that $L_{\rm X}$ is sensitive to $R_{\rm high}$. Very low values of $R_{\rm high}$ are disfavored. $L_{\rm X}$ thus most directly constrains the electron temperature model.

3.6.3 Jet power

Estimates of M87's jet power ($P_{\rm jet}$) have been reviewed in Reynolds et al. (1996); Li et al. (2009); de Gasperin et al. (2012); Broderick et al. (2015); Prieto et al. (2016). The estimates range from 10^{42} to $10^{45}\,{\rm erg\,sec^{-1}}$. This wide range is a consequence of both physical uncertainties in the models used to estimate $P_{\rm jet}$ and the wide range in length- and time- scales probed by the observations. Some estimates may sample a different epoch and thus provide little information on the state of the central engine during EHT2017. Nevertheless, observations of HST-1 yield $P_{\rm jet} \sim 10^{44}\,{\rm erg\,sec^{-1}}$ (e.g. Stawarz et al., 2006). HST-1 is within $\sim 70\,{\rm pc}$ of the central engine and, taking account of relativistic time foreshortening, may be sampling

the central engine $P_{\rm jet}$ over the last few decades. Furthermore, the 1.3mm lightcurve of M87 as observed by SMA shows $\lesssim 50\%$ variability over decade timescales (Bower et al., 2015). Based on these considerations it seems reasonable to adopt a very conservative lower limit on jet power $\equiv P_{\rm jet,min} = 10^{42}\,{\rm erg\,sec^{-1}}$.

To apply this constraint we must define and measure $P_{\rm jet}$ in our models. Our procedure is discussed in detail in § 3.7. In brief, we measure the total energy flux in outflowing regions over the polar caps of the black hole in which the energy per unit rest mass exceeds $2.2c^2$, which corresponds to $\beta\gamma = 1$, where $\beta \equiv v/c$ and γ is Lorentz factor. The effect of changing this cutoff is also discussed in § 3.7. Because the cutoff is somewhat arbitrary, we also calculate $P_{\rm out}$ by including the energy flux in *all* outflowing regions over the polar caps of the black hole; that is, it includes the energy flux in any wide angle, low-velocity wind. $P_{\rm out}$ represents a maximal definition of jet power. Table 3.3 in § 3.7 shows $P_{\rm jet}$ as well as a total outflow power $P_{\rm out}$.

The constraint $P_{\rm jet} > P_{\rm jet,min} = 10^{42}\,{\rm erg\,sec^{-1}}$ rejects all $a_* = 0$ models. This conclusion is not sensitive to the definition of $P_{\rm jet}$: all $a_* = 0$ models also have total outflow power $P_{\rm out} < 10^{42}\,{\rm erg\,sec^{-1}}$. The most powerful $a_* = 0$ model is a MAD model with $R_{\rm high} = 160$, which has $P_{\rm out} = 3.7 \times 10^{41}\,{\rm erg\,sec^{-1}}$ and $P_{\rm jet}$ consistent with 0. We conclude that our $a_* = 0$ models are ruled out.

It is unlikely that the $a_* = 0$ models can be saved by changing the eDF as there is no evidence from the GRMHD simulations that these models are capable of producing a relativistic outflow with $\beta\gamma > 1$. Suppose, however, that we are willing to identify the nonrelativistic outflow, whose power is measured by $P_{\rm out}$, with the jet. Can $P_{\rm out}$ be raised to meet our conservative threshold on jet power? Here the answer is yes, in principle, and this can be done by changing the eDF. The eDF and $P_{\rm out}$ are coupled because $P_{\rm out}$ is determined by $\dot{M}$, and $\dot{M}$ is adjusted to produce the observed compact mm flux. The relationship between $\dot{M}$ and mm flux depends on the eDF. If the eDF is altered to produce mm photons less efficiently (for example, by lowering T_e in a thermal model) then $\dot{M}$ and therefore $P_{\rm out}$

increase. A typical non-thermal eDF, by contrast, is likely to produce mm photons with *greater* efficiency by shifting electrons out of the thermal core and into a non-thermal tail. It will therefore lower $\dot{M}$ and thus $P_{\rm out}$. A thermal eDF with lower T_e could have higher $P_{\rm out}$, as is evident in the large $R_{\rm high}$ SANE models in Table 3.3. There are observational and theoretical lower limits on T_e, however, including a lower limit provided by the observed brightness temperature. As T_e decreases, n_e and B both increase. This has implications for the source linear polarization (Mościbrodzka et al., 2017; Jiménez-Rosales & Dexter, 2018) and is explored in Chapter 4. As T_e declines and n_e and n_i increase, there is also an increase in energy transfer from ions to electrons by Coulomb coupling, and this sets a floor on T_e.

The requirement that $P_{\rm jet} > P_{\rm jet,min}$ eliminates many models other than the $a_* = 0$ models. All SANE models with $|a_*| = 0.5$ fail to produce jets with the required minimum power. Indeed, they also fail the less restrictive condition $P_{\rm out} > P_{\rm jet,min}$, so this conclusion is insensitive to the definition of the jet. We conclude that among the SANE models, only high spin models survive.

At this point it is worth revisiting the SANE, $R_{\rm high} = 1$, $a_* = -0.94$ model that favored a low black hole mass in Section 3.5. These models are not rejected by a naive application of the $P_{\rm jet} > P_{\rm jet,min}$ criterion, but they are marginal. Notice that we needed to assume a mass, however, in applying the this criterion. We have consistently assumed $M = 6.2 \times 10^9 M_\odot$. If we use the $M \sim 3 \times 10^9\, M_\odot$ implied by the best-fit M/D then $\dot{M}$ drops by a factor of two, $P_{\rm jet}$ drops below the threshold, and the model is rejected.

The lower limit on jet power $P_{\rm jet,min} = 10^{42}\,{\rm erg\,sec^{-1}}$ is conservative and the true jet power is likely higher. If we increased $P_{\rm jet,min}$ to $3 \times 10^{42}\,{\rm erg\,sec^{-1}}$, the only surviving models would have $|a_*| = 0.94$ and $R_{\rm high} \geq 10$. This conclusion is also not sensitive to the definition of the jet power: applying the same cut to $P_{\rm out}$ adds only a single model with $|a_*| < 0.94$, the $R_{\rm high} = 160$, $a_* = 0.5$ MAD model. All the remainder have $a_* = 0.94$. Interestingly, the most powerful jets in our ensemble of models are produced by SANE, $a_* = -0.94$, $R_{\rm high} = 160$ models, with $P_{\rm jet} \simeq 10^{43}\,{\rm erg\,sec^{-1}}$.

Estimates for $P_{\rm jet}$ extend to $10^{45}\,{\rm erg\,sec^{-1}}$, but in our ensemble of models the maximum $P_{\rm jet} \sim 10^{43}\,{\rm erg\,sec^{-1}}$. Possible explanations include: (1) $P_{\rm jet}$ is variable and the estimates probe the central engine power at earlier epochs (discussed above), (2) the $P_{\rm jet}$ estimates are too large, or (3) the models are in error. How might our models be modified to produce a larger $P_{\rm jet}$? For a given magnetic field configuration the jet power scales with $\dot{M}c^2$. To increase $P_{\rm jet}$, then, one must reduce the mm flux per accreted nucleon so that at fixed mm flux density $\dot{M}$ increases.[8] Lowering T_e in a thermal model is unlikely to work because lower T_e implies higher synchrotron optical depth, which increases the ring width. We have done a limited series of experiments that suggest that even a modest decrease in T_e would produce a broad ring that is inconsistent with EHT2017 (Event Horizon Telescope Collaboration et al., 2019f). What is required, then, is a nonthermal (or multi-temperature) model with a large population of cold electrons that are invisible at mm wavelength (for a thermal subpopulation, $\Theta_{e,\rm cold} < 1$), and a population of higher energy electrons that produces the observed mm flux (see Falcke & Biermann, 1995).

The $P_{\rm jet}$ in our models is dominated by Poynting flux in the force-free region around the axis (the "funnel"), as in the Blandford & Znajek (1977) force-free magnetosphere model. The energy flux is concentrated along the walls of the funnel.[9] Tchekhovskoy et al. (2011) provided an expression for the energy flux in the funnel, the so-called Blandford–Znajek power $P_{\rm BZ}$, which becomes, in our units,

$$\begin{aligned} P_{\rm BZ} &= 2.8\, f(a_*) \left(\frac{\phi}{15}\right)^2 \dot{M}c^2 \\ &= 2.2\times 10^{43}\ f(a_*) \left(\frac{\phi}{15}\right)^2 \times \\ &\quad \left(\frac{\dot{M}}{10^{-6}\dot{M}_{\rm Edd}}\right)\left(\frac{M}{6.2\times 10^9 M_\odot}\right)\ {\rm erg\,sec^{-1}} \end{aligned} \tag{3.9}$$

[8] The compact mm flux density could be a factor of 2 larger than our assumed 0.5Jy. That would raise $P_{\rm jet}$ by slightly less than a factor of 2.

[9] The total energy flux inside a cone of opening angle θ_0 is proportional to $\sin^4\theta_0$ in the Blandford & Znajek (1977) monopole model if the field strength is fixed, and $\sin^2\theta_0$ if the magnetic flux is fixed.

where $f(a_*) \approx a_*^2(1+\sqrt{1-a_*^2})^{-2}$ (a good approximation for $a_* < 0.95$) and $\dot{M}_{\rm Edd} = 137 M_\odot\,{\rm yr}^{-1}$ for $M = 6.2\times10^9 M_\odot$. This expression was developed for models with a thin disk in the equatorial plane. $P_{\rm BZ}$ is lower for models where the force-free region is excluded by a thicker disk around the equatorial plane. Clearly $P_{\rm BZ}$ is comparable to observational estimates of $P_{\rm jet}$.

In our models (see Table 3.3) $P_{\rm jet}$ follows the above scaling relation but with a smaller coefficient. The ratio of coefficients is model dependent and varies from 0.15 to 0.83. This is likely because the force-free region is restricted to a cone around the poles of the black hole and the width of the cone varies by model. Indeed, the coefficient is larger for MAD than for SANE models, which is consistent with this idea since MAD models have a wide funnel and SANE models have a narrow funnel. This also suggests that future comparison of synthetic 43 and 86 GHz images from our models with lower frequency VLBI data may further constrain the magnetic flux on the black hole.

The connection between the Poynting flux in the funnel and black hole spin has been discussed for some time in the simulation literature, beginning with McKinney & Gammie (2004, see also McKinney 2006; McKinney & Narayan 2007). The structure of the funnel magnetic field can be time-averaged and shown to match the analytic solution of Blandford & Znajek (1977). Furthermore, the energy flux density can be time-averaged and traced back to the event horizon. Is the energy contained in black hole spin sufficient to drive the observed jet over the jet lifetime? The spindown timescale is $\tau = (M - M_{\rm irr})c^2/P_{\rm jet}$ where $M_{\rm irr} \equiv M((1+\sqrt{1-a_*^2})/2)^{1/2}$ is the irreducible mass of the black hole. For the $a_* = 0.94$ MAD model with $R_{\rm high} = 160$, $\tau = 7.3\times10^{12}\,{\rm yr}$, which is long compared to a Hubble time ($\sim10^{10}$ yr). Indeed, the spindown time for all models is long compared to the Hubble time.

We conclude that for models that have sufficiently powerful jets and that are consistent with EHT2017, $P_{\rm jet}$ is driven by extraction of black hole spin energy through the Blandford–Znajek process.

3.6.4 Constraint summary

We have applied constraints from Average Image Scoring, a radiative self-consistency constraint, a constraint on maximum X-ray luminosity, and a constraint on minimum jet power. Which models survive? Here we consider only models for which we have calculated $L_{\rm X}$ and $L_{\rm bol}$ and for which $i = 163^\circ$ (for $a_* \geq 0$) or $i = 17^\circ$ (for $a_* < 0$). Table 3.2 summarizes the results. The first three columns give the model parameters. The next four columns show the result of application of each constraint: THEMIS-AIS (here broken out by individual model rather than groups of models), radiative efficiency ($\epsilon < \epsilon_{\rm thin\,disk}$), $L_{\rm X}$, and $P_{\rm jet}$.

The final column gives the logical **AND** of the previous four columns and allows a model to pass only if it passes all tests. Evidently most of the SANE models fail, with the exception of some $a_* = -0.94$ models and one large $R_{\rm high}$ model with $a_* = 0.94$. A much larger fraction of the MAD models pass, although $a_* = 0$ models all fail because of inadequate jet power. MAD models with small $R_{\rm high}$ also fail. It is the jet power constraint that rejects the largest number of models.

3.7 Simulation results and the $\beta\gamma$ jet cut

Below we provide a table of simulation results for models with a standard inclination of 17° between the approaching jet and the line of sight. In the notation of this paper this corresponds to $i = 17^\circ$ for $a_* < 0$ or $i = 163^\circ$ for $a_* \geq 0$. The table shows models for which we were able to calculate $L_{\rm bol}$ and $L_{\rm X}$. When M is needed to calculate, e.g., $P_{\rm jet}$, we assume $M = 6.2 \times 10^9 M_\odot$.

The first, third, and fourth columns in the table identify the model parameters: SANE or MAD based on dimensionless flux, a_*, and $R_{\rm high}$. Once these parameters are specified, an average value of $\dot{M}$ for the model, which is shown in the last column, can be found from the requirement that the average flux density of 1.3 mm emission is ~ 0.5 Jy (see EHTC IV). This $\dot{M}$ is shown in units of the Eddington accretion rate $\dot{M}_{\rm Edd} = 137 M_\odot\,{\rm yr}^{-1}$.

Flux	a_*	$R_{\rm high}$	AIS	ϵ	$L_{\rm X}$	$P_{\rm jet}$	Result
SANE	-0.94	1	Fail	Pass	Pass	Pass	Fail
SANE	-0.94	10	Pass	Pass	Pass	Pass	**Pass**
SANE	-0.94	20	Pass	Pass	Pass	Pass	**Pass**
SANE	-0.94	40	Pass	Pass	Pass	Pass	**Pass**
SANE	-0.94	80	Pass	Pass	Pass	Pass	**Pass**
SANE	-0.94	160	Fail	Pass	Pass	Pass	Fail
SANE	-0.5	1	Pass	Pass	Fail	Fail	Fail
SANE	-0.5	10	Pass	Pass	Fail	Fail	Fail
SANE	-0.5	20	Pass	Pass	Pass	Fail	Fail
SANE	-0.5	40	Pass	Pass	Pass	Fail	Fail
SANE	-0.5	80	Fail	Pass	Pass	Fail	Fail
SANE	-0.5	160	Pass	Pass	Pass	Fail	Fail
SANE	0	1	Pass	Pass	Pass	Fail	Fail
SANE	0	10	Pass	Pass	Pass	Fail	Fail
SANE	0	20	Pass	Pass	Fail	Fail	Fail
SANE	0	40	Pass	Pass	Pass	Fail	Fail
SANE	0	80	Pass	Pass	Pass	Fail	Fail
SANE	0	160	Pass	Pass	Pass	Fail	Fail
SANE	$+0.5$	1	Pass	Pass	Pass	Fail	Fail
SANE	$+0.5$	10	Pass	Pass	Pass	Fail	Fail
SANE	$+0.5$	20	Pass	Pass	Pass	Fail	Fail
SANE	$+0.5$	40	Pass	Pass	Pass	Fail	Fail
SANE	$+0.5$	80	Pass	Pass	Pass	Fail	Fail
SANE	$+0.5$	160	Pass	Pass	Pass	Fail	Fail
SANE	$+0.94$	1	Pass	Fail	Pass	Fail	Fail
SANE	$+0.94$	10	Pass	Fail	Pass	Fail	Fail

Table 3.2 continued ...

Flux	a_*	$R_{\rm high}$	AIS	ϵ	$L_{\rm X}$	$P_{\rm jet}$	Result
SANE	+0.94	20	Pass	Pass	Pass	Fail	Fail
SANE	+0.94	40	Pass	Pass	Pass	Fail	Fail
SANE	+0.94	80	Pass	Pass	Pass	Fail	Fail
SANE	+0.94	160	Pass	Pass	Pass	Pass	**Pass**
MAD	−0.94	1	Fail	Fail	Pass	Pass	Fail
MAD	−0.94	10	Fail	Pass	Pass	Pass	Fail
MAD	−0.94	20	Fail	Pass	Pass	Pass	Fail
MAD	−0.94	40	Fail	Pass	Pass	Pass	Fail
MAD	−0.94	80	Fail	Pass	Pass	Pass	Fail
MAD	−0.94	160	Fail	Pass	Pass	Pass	Fail
MAD	−0.5	1	Pass	Fail	Pass	Fail	Fail
MAD	−0.5	10	Pass	Pass	Pass	Fail	Fail
MAD	−0.5	20	Pass	Pass	Pass	Fail	Fail
MAD	−0.5	40	Pass	Pass	Pass	Fail	Fail
MAD	−0.5	80	Pass	Pass	Pass	Pass	**Pass**
MAD	−0.5	160	Pass	Pass	Pass	Pass	**Pass**
MAD	0	1	Pass	Fail	Pass	Fail	Fail
MAD	0	10	Pass	Pass	Pass	Fail	Fail
MAD	0	20	Pass	Pass	Pass	Fail	Fail
MAD	0	40	Pass	Pass	Pass	Fail	Fail
MAD	0	80	Pass	Pass	Pass	Fail	Fail
MAD	0	160	Pass	Pass	Pass	Fail	Fail
MAD	+0.5	1	Pass	Fail	Pass	Fail	Fail
MAD	+0.5	10	Pass	Pass	Pass	Fail	Fail
MAD	+0.5	20	Pass	Pass	Pass	Pass	**Pass**
MAD	+0.5	40	Pass	Pass	Pass	Pass	**Pass**
MAD	+0.5	80	Pass	Pass	Pass	Pass	**Pass**

Table 3.2 continued ...

Flux	a_*	$R_{\rm high}$	AIS	ϵ	$L_{\rm X}$	$P_{\rm jet}$	Result
MAD	+0.5	160	Pass	Pass	Pass	Pass	**Pass**
MAD	+0.94	1	Pass	Fail	Fail	Pass	Fail
MAD	+0.94	10	Pass	Fail	Pass	Pass	Fail
MAD	+0.94	20	Pass	Pass	Pass	Pass	**Pass**
MAD	+0.94	40	Pass	Pass	Pass	Pass	**Pass**
MAD	+0.94	80	Pass	Pass	Pass	Pass	**Pass**
MAD	+0.94	160	Pass	Pass	Pass	Pass	**Pass**

Table 3.2: Rejection Table. Average image scoring (THEMIS-AIS) rejects models if $p \leq 0.01$ (see §3.4 and Table 3.1). ϵ is radiative efficiency—models are rejected if ϵ is larger than the corresponding thin disk efficiency (see §3.6.1). Models are rejected base on the X-ray luminosity cut if $L_{\rm X}\rangle 10^{-2\sigma}$ (see §3.6.2). The $P_{\rm jet}$ jet power constraint rejects models if $P_{\rm jet} \leq 10^{42}\,{\rm erg\,sec^{-1}}$ (see §3.6.3).

The measured average dimensionless magnetic flux ϕ is shown in the second column. Notice that ϕ is determined solely from the GRMHD simulation and is independent of the mass scaling $\mathscr{M}$ and the mass M used to fix the flux density. It is also independent of the electron thermodynamics ($R_{\rm high}$).

The fifth column shows the radiative efficiency, which is the bolometric luminosity $L_{\rm bol}$ over $\dot{M}c^2$. Here $L_{\rm bol}$ was found from a relativistic Monte Carlo radiative transport model that includes synchrotron emission, Compton scattering (all orders), and bremsstrahlung. The Monte Carlo calculation makes no approximations in treating the Compton scattering (see Dolence et al., 2009). Bremsstrahlung is negligible in all models.

The sixth column shows predicted X-ray luminosity $L_{\rm X}$ in the $2-10\,{\rm keV}$ band. This was calculated using the same relativistic Monte Carlo radiative transport model as for $L_{\rm bol}$. There are three numbers in this column: the average $\langle L_{\rm X}\rangle$ (left) of the 20 sample spectra used in the calculation, and a maximum and minimum value. The maximum and minimum are

obtained by taking the standard deviation $\sigma(\log_{10} L_{\rm X})$ and setting the maximum (minimum) to $10^{+2\sigma}\langle L_{\rm X}\rangle$ ($10^{-2\sigma}\langle L_{\rm X}\rangle$).

The seventh column shows the jet power

$$P_{\rm jet} \equiv \int_{\beta\gamma>(\beta\gamma)_{\rm cut}} d\theta \, \frac{1}{\Delta t} \int dt d\phi \, \sqrt{-g} \, \left(-T^r{}_t - \rho u^r\right). \tag{3.10}$$

The integral is evaluated at $r = 40\, r_{\rm g}$ for SANE models and $r = 100\, r_{\rm g}$ for MAD models. These radii were chosen because they are close to the outer boundary of the computational domain. Here Δt is the duration of the time-average, $-T^r{}_t$ is a component of the stress–energy tensor representing outward radial energy flux, g is the determinant of the (covariant) metric, ρ is the rest-mass density, and u^r is the radial component of the four-velocity. Here we use Kerr–Schild t, r, θ, ϕ for clarity; in practice the integral is evaluated in simulation coordinates. The quantity in parentheses is the outward energy flux with the rest-mass energy flux subtracted off. The θ integral is done after time averaging and azimuthal integration over the region where

$$(\beta\gamma)^2 \equiv \left(\frac{-T^r{}_t}{\rho u^r}\right)^2 - 1 \;\; > (\beta\gamma)^2_{\rm cut}. \tag{3.11}$$

Here $\beta\gamma$ would be the radial four-velocity as $r \to \infty$ if the flow were steady and all internal magnetic and internal energy were converted to kinetic energy. In Table 3.3, we use $(\beta\gamma)^2_{\rm cut} = 1$ to define the jet. This is equivalent to restricting the jet to regions where the total energy per unit rest mass (including the rest-mass energy) exceeds $\sqrt{5}c^2 \simeq 2.2c^2$.

The ninth column shows the total outflow power $P_{\rm out}$, defined using the same integral as in equation 3.10, but with the θ integral carried out over the entire region around the poles where there is steady outflow (and $\theta < 1$, although the result is insensitive to this condition). $P_{\rm out}$ thus includes both the narrow, fast, relativistic jet and any wide-angle, slow, nonrelativistic outflow. It is the maximal $P_{\rm jet}$ under any definition of jet power.

Finally, the tenth column shows the ratio of the electromagnetic to total energy flux in the jet. In most cases this number is close to 1, i.e., the jet is Poynting dominated. This

measurement is sensitive to the numerical treatment of low density regions in the jet where the jet can be artificially loaded with plasma by numerical floors in the GRMHD evolution. More accurate treatment of the funnel would raise values in this column.

Our choice of $(\beta\gamma)^2_{\rm cut}$, and therefore $P_{\rm jet}$, is somewhat arbitrary. To probe the sensitivity of $P_{\rm jet}$ to $(\beta\gamma)^2_{\rm cut}$, Figure 3.10 shows the ratio $P_{\rm jet}/P_{\rm out}$ (which is determined by the GRMHD model and is thus independent of the electron thermodynamics, i.e., $R_{\rm high}$) as a function of $(\beta\gamma)^2_{\rm cut}$.

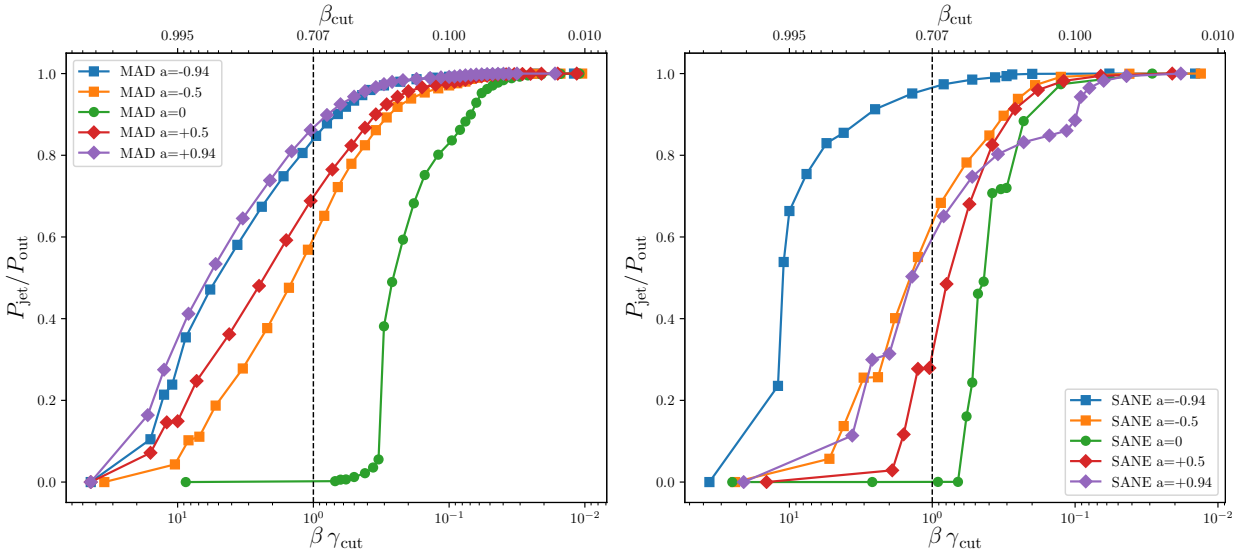


Figure 3.10: Ratio $P_{\rm jet}/P_{\rm out}$ as a function of the outflow velocity cutoff parameter $\beta\gamma_{\rm cut}$. Evidently, as the cut is decreased, so that the maximum asymptotic speed of the jet flow is decreased, an increasing fraction of $P_{\rm out}$ is classified as $P_{\rm jet}$. Our nominal cutoff is $\beta\gamma = 1$, which corresponds to $\beta \equiv v^r/c = 1/\sqrt{2}$. Using this definition, $P_{\rm jet}$ for $a_* = 0$ models is small because the energy flux in the relativistic outflow is small.

The eighth and tenth columns show the jet and outflow efficiency. This is determined by the GRMHD evolution, i.e., it is independent of electron thermodynamics ($R_{\rm high}$). It is > 0.1 only for MAD models with $a_* \geq 0.5$.

The eleventh column shows the fraction of $P_{\rm jet}$ in Poynting flux. This fraction is large for all models, and meaningless for the $a_* = 0$ models which have $P_{\rm jet}$ that is so small that it is difficult to measure accurately.

The problem of defining $P_{\rm jet}$ and $P_{\rm out}$ has been discussed extensively in the literature (see

especially Narayan et al., 2012; Yuan et al., 2015), where alternative definitions based on unbound regions in addition to the jet have been used. Some definitions use the on a fluid Bernoulli parameter $B_e \equiv -u_t(\rho + u + p)/\rho - 1$ while others use μ (the ratio of energy flux to rest mass flux), which is directly related to our $\beta\gamma$.

Flux	ϕ	a_*	$R_{\rm high}$	$L_{\rm bol}/\left(\dot{M}\,c^2\right)$	$L_{\rm X}$ [cgs]	$P_{\rm jet}$ [cgs]	$P_{\rm jet}/\left(\dot{M}\,c^2\right)$	$P_{\rm out}$ [cgs]	$P_{\rm out}/\left(\dot{M}\,c^2\right)$	$P_{\rm jet,em}/P_{\rm jet}$	$\dot{M}/\dot{M}_{\rm Edd}$
SANE	1.02	-0.94	1	1.27×10^{-2}	$3.18^{<49.55}_{>0.20} \times 10^{41}$	1.16×10^{42}	5.34×10^{-3}	1.19×10^{42}	5.48×10^{-3}	0.84	2.77×10^{-5}
SANE	1.02	-0.94	10	1.6×10^{-3}	$9.62^{<64.42}_{>1.44} \times 10^{40}$	4.94×10^{42}	5.34×10^{-3}	5.07×10^{42}	5.48×10^{-3}	0.84	1.19×10^{-4}
SANE	1.02	-0.94	20	6.09×10^{-4}	$3.26^{<11.86}_{>0.90} \times 10^{40}$	5.8×10^{42}	5.34×10^{-3}	5.96×10^{42}	5.48×10^{-3}	0.84	1.39×10^{-4}
SANE	1.02	-0.94	40	2.45×10^{-4}	$8.89^{<50.53}_{>1.56} \times 10^{39}$	7.02×10^{42}	5.34×10^{-3}	7.21×10^{42}	5.48×10^{-3}	0.84	1.69×10^{-4}
SANE	1.02	-0.94	80	1.33×10^{-4}	$2.65^{<18.26}_{>0.39} \times 10^{39}$	8.89×10^{42}	5.34×10^{-3}	9.13×10^{42}	5.48×10^{-3}	0.84	2.13×10^{-4}
SANE	1.02	-0.94	160	7.12×10^{-5}	$6.36^{<55.27}_{>0.73} \times 10^{38}$	1.2×10^{43}	5.34×10^{-3}	1.23×10^{43}	5.48×10^{-3}	0.84	2.87×10^{-4}
SANE	1.11	-0.5	1	1.62×10^{-2}	$1.97^{<3.94}_{>0.98} \times 10^{41}$	2.62×10^{40}	1.86×10^{-4}	3.84×10^{40}	2.72×10^{-4}	0.88	1.81×10^{-5}
SANE	1.11	-0.5	10	2.17×10^{-3}	$1.94^{<5.40}_{>0.69} \times 10^{41}$	1.95×10^{41}	1.86×10^{-4}	2.85×10^{41}	2.72×10^{-4}	0.88	1.34×10^{-4}
SANE	1.11	-0.5	20	6.69×10^{-4}	$3.72^{<7.72}_{>1.80} \times 10^{40}$	2.26×10^{41}	1.86×10^{-4}	3.31×10^{41}	2.72×10^{-4}	0.88	1.56×10^{-4}
SANE	1.11	-0.5	40	2.47×10^{-4}	$9.44^{<13.37}_{>6.67} \times 10^{39}$	2.62×10^{41}	1.86×10^{-4}	3.83×10^{41}	2.72×10^{-4}	0.88	1.81×10^{-4}
SANE	1.11	-0.5	80	1.26×10^{-4}	$1.23^{<4.58}_{>0.33} \times 10^{39}$	3.2×10^{41}	1.86×10^{-4}	4.68×10^{41}	2.72×10^{-4}	0.88	2.21×10^{-4}
SANE	1.11	-0.5	160	7.86×10^{-5}	$3.72^{<16.68}_{>0.83} \times 10^{38}$	4.21×10^{41}	1.86×10^{-4}	6.16×10^{41}	2.72×10^{-4}	0.88	2.9×10^{-4}
SANE	0.99	0	1	3.17×10^{-2}	$2.08^{<194.22}_{>0.02} \times 10^{41}$	2.24×10^{36}	4.4×10^{-8}	5.22×10^{39}	1.03×10^{-4}	1.01	6.5×10^{-6}
SANE	0.99	0	10	1.88×10^{-2}	$4.2^{<425.40}_{>0.04} \times 10^{42}$	4.38×10^{37}	4.4×10^{-8}	1.02×10^{41}	1.03×10^{-4}	1.01	1.27×10^{-4}
SANE	0.99	0	20	5.83×10^{-3}	$1.57^{<39.69}_{>0.06} \times 10^{42}$	8.02×10^{37}	4.4×10^{-8}	1.87×10^{41}	1.03×10^{-4}	1.01	2.33×10^{-4}
SANE	0.99	0	40	7.8×10^{-4}	$8.92^{<41.45}_{>1.92} \times 10^{40}$	9.16×10^{37}	4.4×10^{-8}	2.14×10^{41}	1.03×10^{-4}	1.01	2.66×10^{-4}
SANE	0.99	0	80	1.69×10^{-4}	$2.5^{<19.17}_{>0.33} \times 10^{39}$	1.03×10^{38}	4.4×10^{-8}	2.41×10^{41}	1.03×10^{-4}	1.01	3×10^{-4}
SANE	0.99	0	160	1.08×10^{-4}	$3.44^{<13.32}_{>0.89} \times 10^{38}$	1.23×10^{38}	4.4×10^{-8}	2.87×10^{41}	1.03×10^{-4}	1.01	3.57×10^{-4}
SANE	1.10	0.5	1	4.97×10^{-2}	$5.5^{<34.41}_{>0.88} \times 10^{40}$	2.57×10^{39}	1.63×10^{-4}	9.19×10^{39}	5.86×10^{-4}	0.88	2.01×10^{-6}
SANE	1.10	0.5	10	5.98×10^{-3}	$4.73^{<88.59}_{>0.25} \times 10^{40}$	1.91×10^{40}	1.64×10^{-4}	6.84×10^{40}	5.86×10^{-4}	0.88	1.5×10^{-5}
SANE	1.10	0.5	20	3.33×10^{-3}	$3.83^{<49.18}_{>0.30} \times 10^{40}$	4.09×10^{40}	1.64×10^{-4}	1.47×10^{41}	5.86×10^{-4}	0.88	3.2×10^{-5}
SANE	1.10	0.5	40	1.74×10^{-3}	$2.52^{<22.73}_{>0.28} \times 10^{40}$	8.02×10^{40}	1.64×10^{-4}	2.87×10^{41}	5.86×10^{-4}	0.88	6.28×10^{-5}
SANE	1.10	0.5	80	6.95×10^{-4}	$7.84^{<91.92}_{>0.67} \times 10^{39}$	1.27×10^{41}	1.64×10^{-4}	4.55×10^{41}	5.86×10^{-4}	0.88	9.95×10^{-5}
SANE	1.10	0.5	160	2.78×10^{-4}	$1.37^{<22.85}_{>0.08} \times 10^{39}$	1.69×10^{41}	1.63×10^{-4}	6.06×10^{41}	5.86×10^{-4}	0.88	1.33×10^{-4}
SANE	1.64	0.94	1	1.4	$2.38^{<359.03}_{>0.02} \times 10^{41}$	2.2×10^{40}	7.76×10^{-3}	3.38×10^{40}	1.19×10^{-2}	0.82	3.63×10^{-7}
SANE	1.64	0.94	10	2.7×10^{-1}	$2.79^{<508.99}_{>0.02} \times 10^{41}$	1.4×10^{41}	7.76×10^{-3}	2.15×10^{41}	1.19×10^{-2}	0.82	2.31×10^{-6}

Table 3.3 continued . . .

Flux	ϕ	a_*	$R_{\rm high}$	$L_{\rm bol}/\left(\dot{M}\,c^2\right)$	$L_{\rm X}$ [cgs]	$P_{\rm jet}$ [cgs]	$P_{\rm jet}/\left(\dot{M}\,c^2\right)$	$P_{\rm out}$ [cgs]	$P_{\rm out}/\left(\dot{M}\,c^2\right)$	$P_{\rm jet,em}/P_{\rm jet}$	$\dot{M}/\dot{M}_{\rm Edd}$
SANE	1.64	0.94	20	1.74×10^{-1}	$5.75^{<1685.98}_{>0.02} \times 10^{41}$	3.22×10^{41}	7.76×10^{-3}	4.94×10^{41}	1.19×10^{-2}	0.82	5.31×10^{-6}
SANE	1.64	0.94	40	7.2×10^{-2}	$4.71^{<2490.36}_{>0.01} \times 10^{41}$	5.97×10^{41}	7.76×10^{-3}	9.17×10^{41}	1.19×10^{-2}	0.82	9.84×10^{-6}
SANE	1.64	0.94	80	2.38×10^{-2}	$1.42^{<860.83}_{>0.00} \times 10^{41}$	8.87×10^{41}	7.76×10^{-3}	1.36×10^{42}	1.19×10^{-2}	0.82	1.46×10^{-5}
SANE	1.64	0.94	160	8.45×10^{-3}	$3.22^{<1687.88}_{>0.01} \times 10^{40}$	1.23×10^{42}	7.76×10^{-3}	1.89×10^{42}	1.19×10^{-2}	0.82	2.03×10^{-5}
MAD	8.04	-0.94	1	7.61×10^{-1}	$2.12^{<17.74}_{>0.25} \times 10^{41}$	1.36×10^{42}	2.09×10^{-1}	1.6×10^{42}	2.46×10^{-1}	0.75	8.32×10^{-7}
MAD	8.04	-0.94	10	7.54×10^{-2}	$5.76^{<68.06}_{>0.49} \times 10^{40}$	1.97×10^{42}	2.09×10^{-1}	2.32×10^{42}	2.46×10^{-1}	0.75	1.21×10^{-6}
MAD	8.04	-0.94	20	3.76×10^{-2}	$2.27^{<29.09}_{>0.18} \times 10^{40}$	2.38×10^{42}	2.09×10^{-1}	2.8×10^{42}	2.46×10^{-1}	0.75	1.46×10^{-6}
MAD	8.04	-0.94	40	2.07×10^{-2}	$6.18^{<77.36}_{>0.49} \times 10^{39}$	3×10^{42}	2.09×10^{-1}	3.54×10^{42}	2.46×10^{-1}	0.75	1.84×10^{-6}
MAD	8.04	-0.94	80	1.17×10^{-2}	$1.32^{<26.36}_{>0.07} \times 10^{39}$	3.99×10^{42}	2.09×10^{-1}	4.71×10^{42}	2.46×10^{-1}	0.75	2.45×10^{-6}
MAD	8.04	-0.94	160	6.52×10^{-3}	$2.57^{<46.76}_{>0.14} \times 10^{38}$	5.7×10^{42}	2.09×10^{-1}	6.73×10^{42}	2.46×10^{-1}	0.75	3.5×10^{-6}
MAD	12.25	-0.5	1	2.96×10^{-1}	$1.39^{<11.56}_{>0.17} \times 10^{41}$	3.43×10^{41}	4.91×10^{-2}	6.04×10^{41}	8.64×10^{-2}	0.82	8.95×10^{-7}
MAD	12.25	-0.5	10	4.53×10^{-2}	$2.43^{<19.86}_{>0.30} \times 10^{40}$	5.31×10^{41}	4.92×10^{-2}	9.33×10^{41}	8.64×10^{-2}	0.82	1.38×10^{-6}
MAD	12.25	-0.5	20	2.67×10^{-2}	$8.18^{<77.51}_{>0.86} \times 10^{39}$	6.45×10^{41}	4.92×10^{-2}	1.13×10^{42}	8.64×10^{-2}	0.82	1.68×10^{-6}
MAD	12.25	-0.5	40	1.69×10^{-2}	$2.17^{<22.33}_{>0.21} \times 10^{39}$	8.07×10^{41}	4.92×10^{-2}	1.42×10^{42}	8.64×10^{-2}	0.82	2.1×10^{-6}
MAD	12.25	-0.5	80	1.07×10^{-2}	$4.87^{<50.76}_{>0.47} \times 10^{38}$	1.05×10^{42}	4.92×10^{-2}	1.85×10^{42}	8.64×10^{-2}	0.82	2.74×10^{-6}
MAD	12.25	-0.5	160	6.43×10^{-3}	$1.09^{<7.06}_{>0.17} \times 10^{38}$	1.46×10^{42}	4.92×10^{-2}	2.57×10^{42}	8.64×10^{-2}	0.82	3.81×10^{-6}
MAD	15.44	0	1	2.67×10^{-1}	$1.22^{<14.60}_{>0.10} \times 10^{41}$	0.0	0.0	8.39×10^{40}	1.51×10^{-2}	0.00	7.12×10^{-7}
MAD	15.44	0	10	4.53×10^{-2}	$1.86^{<31.55}_{>0.11} \times 10^{40}$	0.0	0.0	1.39×10^{41}	1.51×10^{-2}	0.00	1.18×10^{-6}
MAD	15.44	0	20	2.81×10^{-2}	$5.98^{<101.81}_{>0.35} \times 10^{39}$	0.0	0.0	1.71×10^{41}	1.51×10^{-2}	0.00	1.46×10^{-6}
MAD	15.44	0	40	1.85×10^{-2}	$1.63^{<27.75}_{>0.10} \times 10^{39}$	0.0	0.0	2.15×10^{41}	1.51×10^{-2}	0.00	1.82×10^{-6}
MAD	15.44	0	80	1.21×10^{-2}	$3.51^{<61.34}_{>0.20} \times 10^{38}$	0.0	0.0	2.77×10^{41}	1.51×10^{-2}	0.00	2.35×10^{-6}
MAD	15.44	0	160	7.63×10^{-3}	$8.06^{<80.62}_{>0.81} \times 10^{37}$	0.0	0.0	3.73×10^{41}	1.51×10^{-2}	0.00	3.17×10^{-6}
MAD	15.95	0.5	1	5.45×10^{-1}	$1.57^{<11.98}_{>0.21} \times 10^{41}$	4.64×10^{41}	1.16×10^{-1}	6.74×10^{41}	1.69×10^{-1}	0.85	5.11×10^{-7}
MAD	15.95	0.5	10	9.45×10^{-2}	$2.71^{<36.30}_{>0.20} \times 10^{40}$	8.07×10^{41}	1.16×10^{-1}	1.17×10^{42}	1.69×10^{-1}	0.85	8.89×10^{-7}
MAD	15.95	0.5	20	5.54×10^{-2}	$9.67^{<126.69}_{>0.74} \times 10^{39}$	1.02×10^{42}	1.16×10^{-1}	1.49×10^{42}	1.69×10^{-1}	0.85	1.13×10^{-6}
MAD	15.95	0.5	40	3.5×10^{-2}	$3.3^{<39.01}_{>0.28} \times 10^{39}$	1.32×10^{42}	1.16×10^{-1}	1.92×10^{42}	1.69×10^{-1}	0.85	1.45×10^{-6}

Table 3.3 continued . . .

Flux	ϕ	a_*	$R_{\rm high}$	$L_{\rm bol}/\left(\dot{M}\,c^2\right)$	$L_{\rm X}$ [cgs]	$P_{\rm jet}$ [cgs]	$P_{\rm jet}/\left(\dot{M}\,c^2\right)$	$P_{\rm out}$ [cgs]	$P_{\rm out}/\left(\dot{M}\,c^2\right)$	$P_{\rm jet,em}/P_{\rm jet}$	$\dot{M}/\dot{M}_{\rm Edd}$
MAD	15.95	0.5	80	2.22×10^{-2}	$8^{<91.84}_{>0.70} \times 10^{38}$	1.74×10^{42}	1.16×10^{-1}	2.52×10^{42}	1.69×10^{-1}	0.85	1.92×10^{-6}
MAD	15.95	0.5	160	1.35×10^{-2}	$1.79^{<8.44}_{>0.38} \times 10^{38}$	2.38×10^{42}	1.16×10^{-1}	3.46×10^{42}	1.69×10^{-1}	0.85	2.62×10^{-6}
MAD	12.78	0.94	1	3.65	$5.19^{<43.60}_{>0.62} \times 10^{41}$	1.97×10^{42}	8.23×10^{-1}	2.29×10^{42}	9.55×10^{-1}	0.80	3.07×10^{-7}
MAD	12.78	0.94	10	3.68×10^{-1}	$1.3^{<13.22}_{>0.13} \times 10^{41}$	3.04×10^{42}	8.23×10^{-1}	3.52×10^{42}	9.55×10^{-1}	0.80	4.73×10^{-7}
MAD	12.78	0.94	20	1.79×10^{-1}	$5^{<56.22}_{>0.44} \times 10^{40}$	3.73×10^{42}	8.23×10^{-1}	4.33×10^{42}	9.55×10^{-1}	0.80	5.81×10^{-7}
MAD	12.78	0.94	40	9.43×10^{-2}	$1.54^{<22.13}_{>0.11} \times 10^{40}$	4.74×10^{42}	8.23×10^{-1}	5.5×10^{42}	9.55×10^{-1}	0.80	7.38×10^{-7}
MAD	12.78	0.94	80	5.19×10^{-2}	$3.74^{<80.85}_{>0.17} \times 10^{39}$	6.26×10^{42}	8.23×10^{-1}	7.27×10^{42}	9.55×10^{-1}	0.80	9.75×10^{-7}
MAD	12.78	0.94	160	2.82×10^{-2}	$6.97^{<186.48}_{>0.26} \times 10^{38}$	8.75×10^{42}	8.23×10^{-1}	1.02×10^{43}	9.55×10^{-1}	0.80	1.36×10^{-6}

Table 3.3: Summary of EHT model comparison metrics for total intensity analysis in Event Horizon Telescope Collaboration et al. (2019e).

3.8 Discussion

We have interpreted the EHT2017 data using a limited library of models with attendant limitations. Many of the limitations stem from the GRMHD model, which treats the plasma as an ideal fluid governed by equations that encode conservation laws for particle number, momentum, and energy. The eDF, in particular, is described by a number density and temperature, rather than a full distribution function, and the electron temperature T_e is assumed to be a function of the local ion temperature and plasma $\beta_{\rm p}$. Furthermore, all models assume a Kerr black hole spacetime, although there are alternatives. Here we consider some of the model limitations and possible extensions.

3.8.1 Radiative effects

Some post-processed GRMHD simulations that are consistent with EHT data and the 1.3 mm flux density have unphysically large radiative efficiencies (see Section 3.6). This implies the radiative cooling timescale is comparable to or less than the advection timescale. As a consequence, including radiative cooling in simulations may be necessary to recover self-consistent models (see Mościbrodzka et al., 2011; Dibi et al., 2012). In our models we use a single parameter $R_{\rm high}$ to adjust T_e and account for all effects that might influence the electron energy density. How good is this approximation?

The importance of radiative cooling can be assessed using newly developed, state-of-the-art general relativistic radiation GRMHD codes. Sądowski et al. (2013b, see also Sądowski et al. 2014; McKinney et al. 2014; Sądowski et al. 2017) applied the M1 closure (Levermore, 1984), which treats the radiation as a relativistic fluid. Ryan et al. (2015) introduced a Monte Carlo radiation GRMHD method, allowing for full frequency dependent radiation transport. Models for turbulent dissipation into the electrons and ions, as well as heating and cooling physics that sets the temperature ratio T_i/T_e, have been added to GRMHD and radiative GRMHD codes and used in simulations of Sgr A* (Ressler et al., 2015, 2017; Chael

et al., 2018a) and M87 (Ryan et al., 2018; Chael et al., 2019). While the radiative cooling and Coulomb coupling physics in these simulations is well understood, the particle heating process, especially the relative heating rates of ions and electrons, remains uncertain.

Radiation GRMHD models are computationally expensive and do not have the same scaling freedom as the GRMHD models, so they need to be repeatedly re-run with different initial conditions until they produce the correct 1.3 mm flux density. It is therefore impractical to survey the parameter space using radiation GRMHD at the present time. It is possible, however, to check individual GRMHD models against existing radiation GRMHD models of M87 (Ryan et al., 2018; Chael et al., 2019).

The SANE radiation GRMHD models of Ryan et al. (2018) with $a_* = 0.94$ and $M = 6 \times 10^9 M_\odot$ can be compared to GRMHD SANE $a_* = 0.94$ models at various values of $R_{\rm high}$. The radiative models have $\dot{M}/\dot{M}_{\rm Edd} = 5.2 \times 10^{-6}$ and $P_{\rm jet} = 5.1 \times 10^{41}\,{\rm erg\,sec^{-1}}$. The GRMHD models in this work have, for $1 \leq R_{\rm high} \leq 160$, $0.36 \times 10^{-6} \leq \dot{M}/\dot{M}_{\rm Edd} \leq 20 \times 10^{-6}$ and $0.22 \leq P_{\rm jet}/(10^{41}\,{\rm erg\,sec^{-1}}) \leq 12$ (Table 3.3). Evidently the mass accretion rates and jet powers in the GRMHD models span a wide range that depends on $R_{\rm high}$, but when we choose $R_{\rm high} = 10 - 20$ they are similar to what is found in the radiative GRMHD model when using the turbulent electron heating model (Howes, 2010).

We have also directly compared the T_e distribution in the emitting region and found that the radiation GRMHD model is quite close to the $R_{\rm high} = 10$ model. The resulting images are qualitatively similar, with an asymmetric photon ring that is brighter in the South and a weak inner ring associated with the funnel wall emission as in Figure 3.2. The radiation GRMHD SANE model, like the vast majority of our nonradiative GRMHD SANE models, would be ruled out by the condition $P_{\rm jet} > 10^{42}\,{\rm erg\,sec^{-1}}$.

The MAD radiation GRMHD models of Chael et al. (2019) with $a_* = 0.94$ and $M = 6.2 \times 10^9 M_\odot$ can be compared to GRMHD MAD $a_* = 0.94$ models at various values of $R_{\rm high}$. Chael et al. (2019) uses two dissipation models: the Howes (2010, hereafter H10) model of heating from a Landau-damped turbulent cascade, and the Rowan et al. (2017, hereafter

R17) model of heating based on simulations of trans-relativistic magnetic reconnection. The (H10, R17) models have $\dot{M}/\dot{M}_{\rm Edd} = (3.6, 2.3) \times 10^{-6}$ and $P_{\rm jet} = (6.6, 13) \times 10^{42}\,{\rm erg\,sec^{-1}}$. The GRMHD models have, for $1 \leq R_{\rm high} \leq 160$, $0.13 \times 10^{-6} \leq \dot{M}/\dot{M}_{\rm Edd} \leq 1.4 \times 10^{-6}$ and $2.3 \leq P_{\rm jet}/(10^{42}\,{\rm erg\,sec^{-1}}) \leq 8.8$ (Table 3.3). In the radiation GRMHD MAD models, $\dot{M}$ lies in the middle of the range spanned by the nonradiative GRMHD models and jet power lies at the upper end of the range spanned by the nonradiative GRMHD models. The T_e distributions in the radiative and nonradiative MAD models differ: the mode of the radiation GRMHD model T_e distribution is about a factor of 3 below the mode of the T_e distribution in the $R_{\rm high} = 20$ GRMHD model, and the GRMHD model has many more zones at $\Theta_e \sim 100$ that contribute to the final image than the radiation GRMHD models. This difference is a consequence of the $R_{\rm high}$ model for T_e: in MAD models almost all the emission emerges at $\beta_{\rm p} \lesssim 1$, so $R_{\rm high}$, which changes T_e in the $\beta_{\rm p} > 1$ region, offers little control over T_e in the emission region. Nevertheless the jet power and accretion rates are similar in the radiative and nonradiative MAD models, and the time-averaged radiative and non-radiative images are qualitatively indistinguishable. This suggests that the image is determined mainly by the spacetime geometry and is insensitive to the details of the plasma evolution.

This review of radiative effects is encouraging but incomplete: it only considers a limited selection of models and a narrow set of observational constraints. Future studies of time-dependence and polarization are likely to sharpen the contrast between radiative and nonradiative models.

3.8.2 Non-thermal electrons

Throughout this work we have considered only a thermal eDF. While a thermal eDF can account for the observed emission at mm wavelengths in M87 (e.g., Prieto et al., 2016; Mościbrodzka et al., 2016; Ryan et al., 2018; Chael et al., 2019), eDFs that include a non-thermal tail can also explain the observed SED (Broderick & Loeb, 2009; Yu et al., 2010; Dexter et al., 2012; Li et al., 2016; Davelaar et al., 2018).

The role of non-thermal electrons (and positrons) in producing the observed compact emission is not a settled question, and it cannot be settled in this first investigation of EHT2017 models; however, there are constraints. The number density, mean velocity, and energy density of the eDF are fixed or limited by the GRMHD models. In addition, the eDF cannot on average sustain features that would be erased by kinetic instabilities on timescales short compared to $r_{\rm g}c^{-1}$. Some non-thermal eDFs increase $F_\nu/\dot{M}$ in comparison to a thermal eDF, implying lower values of $\dot{M}$ than quoted above (Ball et al., 2018; Davelaar et al., 2019). These lower values of $\dot{M}$ can slightly change the source morphology, e.g., by decreasing the visibility of the approaching jet (e.g., Dexter et al., 2012).

One can evaluate the influence of non-thermal eDFs in several ways. For example, it is possible to study simplified, phenomenological models. Emission features due to the cooling of non-thermal electrons may then reveal how and where the non-thermal electrons are produced (Pu et al., 2017). The effect of non-thermal eDFs can also be studied by post-processing of ideal GRMHD models if one assumes that the electrons have a fixed, parameterized form such as a power-law distribution (Dexter et al., 2012) or a κ-distribution (Davelaar et al., 2018). These parameterized models produce SEDs that agree with radio to near-infrared data but they are approximations to the underlying physics and do not resolve the microscopic processes that accelerate particles. One can also include dissipative processes explicitly in the GRMHD models, including scalar resistivity (Palenzuela et al., 2009; Dionysopoulou et al., 2013; Del Zanna et al., 2016; Qian et al., 2016; Ripperda et al., 2019), heat fluxes and viscosities (pressure anisotropies Chandra et al., 2015; Ressler et al., 2015; Foucart et al., 2017), and particle acceleration (e.g., Chael et al., 2017). Ultimately special and general relativistic particle-in-cell codes (Watson & Nishikawa, 2010; Levinson & Cerutti, 2018; Chen et al., 2018; Parfrey et al., 2019) will enable direct investigations of kinetic processes.

3.8.3 Other models and analysis limitations

We have used a number of other approximations in generating our models. Among the most serious are:

(1) Fast Light approximation. A GRMHD simulation produces a set of dump files containing the model state at a single global (Kerr–Schild) coordinate time. Because the dynamical time is only slightly longer than the light-crossing time, one in principle needs to trace rays through a range of coordinate times, i.e., by interpolation between multiple closely spaced dump files. This is difficult in practice because a high cadence of output files is required, limiting the speed of the GRMHD simulations and requiring prohibitively large data storage. In addition, the cost of ray tracing through multiple output files is high. Because of this, we adopt the commonly used fast light approximation in which GRMHD variables are read from a single dump file and held steady during the ray tracing. Including light travel time delays produces minor changes to the small-scale image structure and to light curves (e.g., Dexter et al., 2010; Bronzwaer et al., 2018; Younsi et al., 2016), although it is essential for the study of variability on the light-crossing timescale.

(2) Untilted Disks. We have assumed that the disk angular momentum vector and black hole spin vector are (anti-)aligned. There is no reason for the angular momentum vector of the accretion flow on large scales to align with the black hole spin vector, and there is abundant evidence for misaligned disks in AGN (e.g., Miyoshi et al., 1995). How might disk tilt affect our results? Tilting the disk by as little as $\sim 15^\circ$ is enough to set up a standing, two-armed spiral shock close to the ISCO (Fragile & Blaes, 2008). This shock directly affects the morphology of mm wavelength images, especially at low inclination, in models of Sgr A* (Dexter & Fragile, 2013, , especially Figure 5), producing an obvious two-armed spiral pattern on the sky. If this structure were also present in images of tilted models of M87 then it is possible that even a modest tilt could be ruled out.

If modest tilt is present in M87 it is unlikely to affect our conclusion regarding the sign of black hole spin. That conclusion depends on emission from funnel wall plasma in counter-

rotating ($a_* < 0$) disks. The funnel wall plasma is loaded onto funnel plasma field lines by local instabilities at the wall and then rotates with the funnel and therefore the black hole (Wong et al., 2021a). The funnel wall is already unsteady, fluctuating by tens of degrees in azimuth and in time, so a modest tilt seems unlikely to dramatically alter the funnel wall structure.

Is there observational evidence for tilt in M87? In numerical studies of tilted disks the jet emerges perpendicular to the disk Liska et al. (2018), and tilted disks are expected to precess. One might then expect that a tilted source would produce a jet that exhibits periodic variations, or periodic changes in jet direction with distance from the source, as seen in other sources. There is little evidence of this in M87 (see Park et al., 2019, for a discussion of possible misaligned structure in the jet). Indeed, Walker et al. (2018) sees at most small displacements of the jet with time and distance from the source at mas scales. In sum, there is therefore little observational motivation for considering tilted disk models.

Tilted disk models of M87 are an interesting area for future study. It is possible that the inner disk may align with the black hole via a thick-disk variant of the Bardeen & Petterson (1975) effect. Existing tilted thick disk GRMHD simulations (e.g., Fragile et al., 2007; Shiokawa, 2013; McKinney et al., 2013; Liska et al., 2018) show some evidence for alignment and precession, but understanding of the precession and alignment timescales is incomplete. It will be challenging to extend the Image Library to include a survey of tilted disk models, however, since with tilted disks there are two new parameters: the two angles that describe the orientation of the outer disk with respect to the black hole spin vector and the line of sight.

(3) Pair Production. In some models of M87 the mm emission is dominated by electron-positron pairs within the funnel, even close to the horizon scale (see Beskin et al., 1992; Mościbrodzka et al., 2011; Broderick & Tchekhovskoy, 2015; Levinson & Rieger, 2011; Hirotani & Pu, 2016; Wong et al., 2021b). The pairs are produced from the background radiation field or from a pair cascade process following particle acceleration by unscreened electric

fields, which we cannot evaluate using ideal GRMHD models. We leave it to future work to assess whether these models can plausibly suppress emission from the disk and funnel wall, and simultaneously produce a sufficiently powerful jet.

(4) Numerical Treatment of Low Density Regions. Virtually all MHD simulations, including ours, use a "floor" procedure that resets the density if it falls below a minimum value. If this is not done then truncation error accumulates dramatically in the low density regions and the solution is corrupted. If the volume where floors are activated contains only a small fraction of the simulation mass, momentum, and energy then most aspects of the solution are unaffected by this procedure (e.g. McKinney & Gammie, 2004).

The simulation-reported temperature of the plasma is no longer reliable in regions where the floors are activated. Since floors are commonly activated in regions with $B^2/\rho > 1$, we universally cut off emission in those regions. In models where floors are only activated in the funnel (e.g., most SANE models), the resulting images are insensitive to the choice of cutoff B^2/ρ. In MAD models the regions of low and high density are mixed because lightly loaded magnetic field lines that are trapped in the hole bubbles outward through the disk. In this case emission at $\nu > 230$GHz can be sensitive to the choice of cutoff B^2/ρ Chael et al. (2019). The sense of the effect is that greater cutoff B^2/ρ implies more emission at high frequency. Our use of a cutoff $B^2/\rho = 1$ is therefore likely to underestimate mm emission and therefore overestimate $\dot{M}$ and $P_{\rm jet}$. Accurate treatment of the dynamics and thermodynamics of low density regions and especially sharp boundaries between low and high density regions is a fundamental numerical problem in black hole accretion flow modeling that merits further attention.

3.9 Conclusion

In this paper we have made a first attempt at understanding the physical implications of a single, high quality EHT dataset for M87. We have compared the data to a library of mock

images produced from GRMHD simulations by GRRT calculations. The library covers a parameter space that is substantially larger than earlier model surveys. The results of this comparison are consistent with the hypothesis that the compact 1.3 mm emission in M87 arises within a few $r_{\rm g}$ of a Kerr black hole and that the ringlike structure of the image is generated by strong gravitational lensing and Doppler beaming. The models predict that the asymmetry of the image depends on the sense of black hole spin. If the models are sufficiently representative, then the spin vector of the black hole in M87 points away from Earth (the black hole spins clockwise on the sky). The models also predict that there is a strong energy flux directed away from the poles of the black hole and that this energy flux is electromagnetically dominated. If the models are correct then the central engine for the M87 jet is powered by the electromagnetic extraction of free energy associated with black hole spin via the Blandford–Znajek process.

In our models, M87's compact mm emission is generated by the synchrotron mechanism. Our ability to make physical inferences based on the models is therefore intimately tied to the quality of our understanding of the eDF. We have used a thermal model with a single free parameter that adjusts the ratio of ion to electron temperature in regions with plasma $\beta_{\rm p} > 1$ (i.e., regions where magnetic pressure is less than gas pressure). This simple model does not span the range of possible plasma behavior. The theory of high temperature, collisionless plasmas must be better understood if this core physical uncertainty of sub-Eddington black hole accretion is to be eliminated. At present our understanding is inadequate, and alternative eDF models occupy a large, difficult-to-explore parameter space with the potential to surprise. Despite these uncertainties many of the models produce images with similar morphology that is consistent with EHT2017 data. This suggests that the image shape is controlled mainly by gravitational lensing and the spacetime geometry rather than details of the plasma physics.

Although the EHT2017 images are consistent with the vast majority of our models, parts of the parameter space can be rejected on physical grounds or by comparison with

contemporaneous data at other wavelengths. We reject some models because, even though all models are variable, some models are *too variable* to be consistent with the data. We can also reject models based on a radiative efficiency cut (the models are not self-consistent and would cool quickly if radiative effects were included), an X-ray luminosity cut using contemporaneous *Chandra* and *NuSTAR* data, and on a jet power cut. The requirement that the jet power exceed a conservative lower limit of $10^{42}\,\mathrm{erg\,sec^{-1}}$ turns out to eliminate many models, including all models with $a_* = 0$.

We have examined the astrophysical implications of only a subset of EHT2017 data; much remains to be done, and there are significant opportunities for further constraining the models. EHT2017 data includes tracks from four separate days of observing; each day is $2.8\,r_{\rm g}c^{-1}$ (see EHTC IV). This timescale is short compared to the decorrelation timescale of simulated images, which is $\sim 50\,r_{\rm g}c^{-1}$, and smaller than the light crossing time of the source plasma. Analysis techniques that use short-timescale variations in the data will need to be developed and are likely to recover new, more stringent constraints on the model from the EHT2017 dataset. Comparison of model polarization maps of the source with EHT2017 data sharply limit the space of allowed models (Mościbrodzka et al., 2017) and are considered in Chapter 4. Finally, in this work the only multiwavelength companion data we consider are X-ray observations. Simultaneous data are available at many other wavelengths, from the radio to the gamma-rays. These data are likely to further limit the range of acceptable models and guide the implementation of predictive electron physics models.

We have also not yet considered how the physical properties of the jet are constrained by lower frequency VLBI observations, which constrain jet kinematics (Mertens et al., 2016; Hada et al., 2017; Britzen et al., 2017; Walker et al., 2018; Kim et al., 2018), the jet width profile (Asada & Nakamura, 2012; Hada et al., 2013; Nakamura et al., 2018), the total jet power at kilo-parsec scale (Owen et al., 2000; Stawarz et al., 2006), the jet power (e.g., Kino et al., 2014, 2015), the core shift (Hada et al., 2011), and the symmetric limb-brightening structure (Takahashi et al., 2018; Kim et al., 2018). The jet width profile is potentially very

interesting because it depends on the magnetic flux ϕ: the jet internal magnetic pressure $\propto \phi^2$. We therefore expect (and see in our numerical simulations; see Figure 3.5) that MAD jets are wider at the base than SANE jets.

A second epoch of observations ($\gtrsim 50\, r_{\rm g} c^{-1} \sim 2$ weeks after EHT2017, when the models suggest that source structure will decorrelate) will increase the power of the average image analysis to reject models. The EHT2017 data were able to reject one entire category of models with confidence: high magnetic flux (MAD), retrograde, high spin models. Other categories of models, such as the low magnetic flux, high spin models, are assigned comparatively low probabilities by the average image scoring scheme. Data taken later, more than a decorrelation time after EHT2017 (model decorrelation times are of order two weeks), will provide an independent realization of the source. The probabilities attached to individual models by average image scoring will then multiply. For example, a model with probability 0.05 that is assigned probability 0.05 in comparison to a second epoch of observation would then have probability $0.05^2 = 2.5 \times 10^{-3}$, and would be strongly disfavored by the average image scoring criterion (see Section 3.4).

Future EHT 345 GHz campaigns (EHTC II) will provide excellent constraints, particularly on the width of the ring. The optical depth on every line of sight through the source is expected to decrease (the drop is model and location dependent). In our models this makes the ring narrower, better defined, easier to measure accurately from VLBI data, and less dependent on details of the source plasma model.

Certain features of the model are geometric and should be present in future EHT observations. The photon ring is a persistent feature of the model related to the mass and distance to the black hole. It should be present in the next EHT campaign unless there is a dramatic change in $\dot{M}$, which would be evident in the SED. The asymmetry in the photon ring is also a persistent feature of the model because, we have argued, it is controlled by the black hole spin. The asymmetry should therefore remain in the southern half of the ring for the next EHT campaign, unless there is a dramatic tilt of the inner accretion flow. If the small

scale and large scale jet are aligned, then EHT2017 saw the brightest region at unusually small PA, and future campaigns are likely (but not certain) to see the peak brightness shift further to the West. Future 230 GHz EHT campaigns (EHTC II) will thus sharply test the GRMHD source models.

Together with complementary studies that are presently targeting either the supermassive black-hole candidate at the Galactic Center (Eckart & Genzel, 1997; Ghez et al., 1998; Abuter et al., 2018; Gravity Collaboration et al., 2018) or stellar-mass binary black holes whose gravitational-wave emission is recorded by the LIGO and Virgo detectors (Abbott et al., 2016a), the results provided here are consistent with the existence of astrophysical black holes. More importantly, they clearly indicate that their phenomenology, despite being observed on mass scales that differ by eight orders of magnitude, follows very closely the one predicted by general relativity. This demonstrates the complementarity of experiments studying black holes on all scales, promising much improved tests of gravity in its most extreme regimes.

Chapter 4

Interpreting M87: Linear polarization

In April 2017, the Event Horizon Telescope observed the shadow of the supermassive black hole at the core of the elliptical galaxy Messier 87 at 230 GHz. This chapter covers my contributions to the EHT analysis of linear polarization data from the observation. At the beginning of the chapter, I describe, motivate, and test a modal image decomposition procedure that can be used to quantitatively differentiate between the SANE and MAD accretion states using data from the linear polarization field of the image. Moving on to real-world linear polarization data of the M87 source, I describe a comparison between theoretical models and the observed polarization metrics produced by the EHT and ALMA. The observational polarization data appear to be scrambled on scales smaller than the EHT beam, implying the potential importance of Faraday rotation effects. Through a quantitative comparison with a large library of simulated polarimetric images from GRMHD simulations, it is seem that several physical models can explain critical features of the polarimetric EHT observations while producing the observed jet power from M87. All consistent GRMHD models are of MAD flows, where magnetic fields near the event horizon are dynamically important.

4.1 Introduction

The Event Horizon Telescope (EHT) collaboration recently published total intensity images of event-horizon scale emission around the supermassive black hole in the core of the Messier

This chapter has been adapted from work with D. C. M. Palumbo, B. S. Prather, and collaborators in the Event Horizon collaboration that was published as ApJ Volume 894, Issue 2, id.156, 2020 and ApJL, Volume 910, Issue 1, id.L13 2021.

87 galaxy (hereafter the black hole is called M87, Event Horizon Telescope Collaboration et al. 2019a,b,c,d, hereafter EHTC I; EHTC II; EHTC III; EHTC IV). The data reveal a $42 \pm 3\,\mu$as diameter ring-like structure broadly consistent with the shadow of a black hole as predicted by Einstein's Theory of General Relativity (Event Horizon Telescope Collaboration et al. 2019e,f, hereafter EHTC V; EHTC VI). The brightness temperature of the ring at 230 GHz ($\gtrsim 10^{10}$ K) is naturally explained by synchrotron emission from relativistic electrons gyrating around magnetic field lines. The ring brightness asymmetry results from light bending and Doppler beaming due to relativistic rotation of the matter around the black hole.

The M87 galaxy is best known for launching a kpc-scale FR-I type relativistic jet, whose kinetic power is estimated to be $\sim 10^{42-44}\,\mathrm{erg\,s^{-1}}$ (e.g., Stawarz et al. 2006, de Gasperin et al. 2012). The structure of the relativistic jet has been resolved and studied in great detail at multiple wavelengths from radio to X-rays (e.g., Di Matteo et al. 2003; Harris et al. 2009; Walker et al. 2018; Kim et al. 2018). It is not yet understood exactly how the jet is formed and connected to the central object.

The published EHT image of M87 together with multi-wavelength observations are consistent with the picture that the supermassive black hole in M87 is surrounded by a relativistically hot, magnetized plasma (Rees et al., 1982; Narayan & Yi, 1995a; Narayan et al., 1995; Yuan & Narayan, 2014; Reynolds et al., 1996; Yuan et al., 2002; Di Matteo et al., 2003). However, it is not clear whether the compact ring emission is produced by the plasma that is inflowing (in a thick accretion flow), outflowing (at the jet base or in a wind), or both. Furthermore, the total intensity EHT observations could not constrain the structure of magnetic fields in the observed emission region. Additional information is needed to constrain the detailed physical scenario in M87.

Event Horizon Telescope Collaboration et al. (2021a) (hereafter EHTC VII) reported new polarimetric EHT 2017 observations of M87, which are reproduced in Figure 4.1. These images reveal that a significant fraction of the ring emission is linearly polarized, as expected

for synchrotron radiation. The EHT polarimetric measurements are consistent with unresolved observations of the radio core at the same frequency with the SMA (Kuo et al., 2014) and ALMA (Goddi et al., 2021). They also provide a detailed view of the polarized emission region on event-horizon scales near the black hole.

Polarized synchrotron radiation traces the underlying magnetic field configuration and magnetized plasma properties along the line of sight (Bromley et al., 2001; Broderick & Loeb, 2009; Mościbrodzka et al., 2017). The polarimetric measurements allow us to carry out new *quantitative tests* of horizon scale scenarios for accretion and jet launching around the M87 black hole. In this paper we present our interpretation of the EHTC VII resolved polarimetric images of the ring in M87.

We begin with an introduction to image linear polarization structure in Section 4.2. In Section 4.3 we motivate and define a decomposition procedure for linear polarization data in black hole images. In Section 4.4 we apply the decomposition to a large set of simulated images and detail observed statistical trends, and we discuss model limitations for the decomposition in Section 4.5.

We then report polarimetric constraints from M87 EHT 2017 as well as supplemental observations in Section 4.6. In Section 4.7 we present one-zone estimates of the properties of the synchrotron emitting plasma. Since the M87 core occupies the trans-relativistic regime, we realize a full calculation of polarized radiative transfer using numerical simulations; in Section 4.8, we describe the numerical simulation procedure used to produce accretion flow simulations that can be compared to the observables. In Section 4.9 we show that a limited subset of the simulation parameter space is consistent with the observables, and show that the favored simulations feature dynamically important magnetic fields. Section 4.10 provides an itemized report of model theory metrics. We discuss limitations of our models in Section 4.11 and discuss how future EHT observations can further constrain the magnetic field structure and plasma properties near the supermassive black hole's event horizon in Section 4.12.

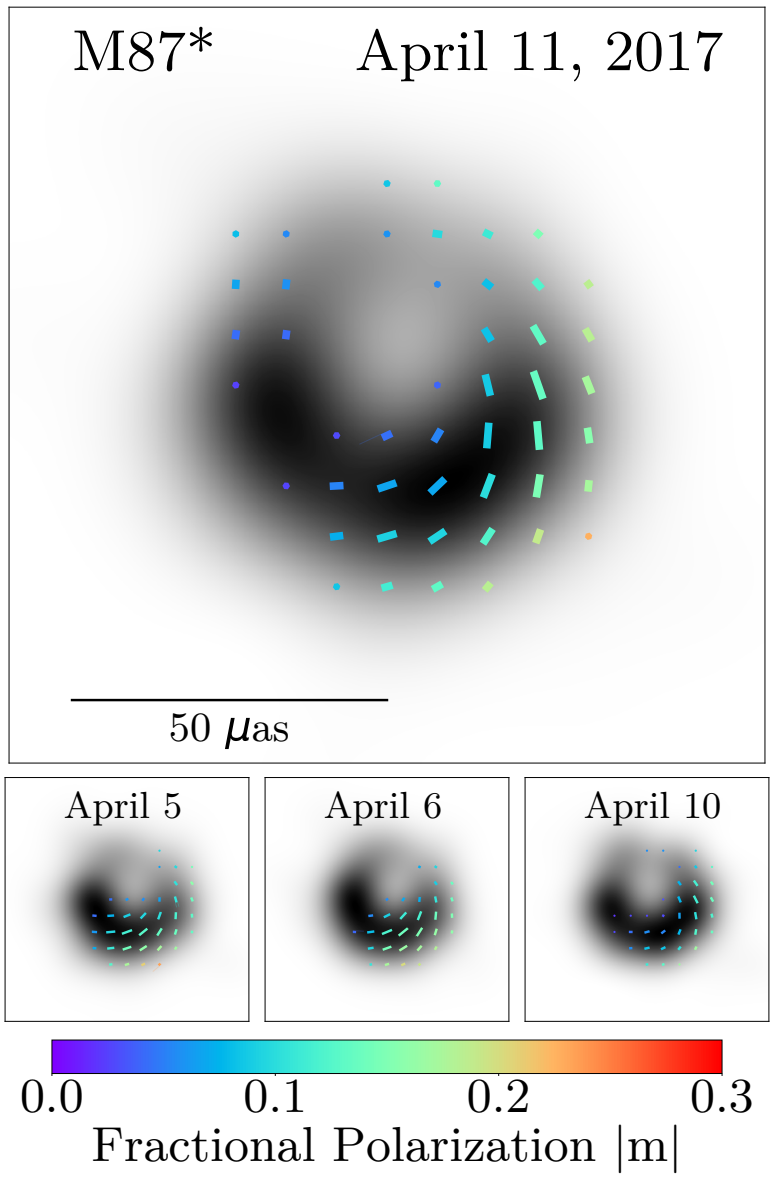


Figure 4.1: *Top*: April 11 fiducial polarimetric image of M87 from EHTC VII. The gray scale encodes the total intensity, and ticks illustrate the degree and direction of linear polarization. The tick color indicates the amplitude of the fractional linear polarization, the tick length is proportional to $|\mathcal{P}| \equiv \sqrt{\mathcal{Q}^2 + \mathcal{U}^2}$, and the tick direction indicates the EVPA, or electric-vector linear polarization angle. Polarization ticks are displayed only in regions where $\mathcal{I} > 10\% \mathcal{I}_{\rm max}$ and $|\mathcal{P}| > 20\% |\mathcal{P}|_{\rm max}$. *Bottom:* Polarimetric images of M87 taken on different days.

4.2 Azimuthal structure in black hole images

The accreting material around the M87 supermassive black hole is typically modeled as a radiatively inefficient accretion flow forming a geometrically thick disk of infalling plasma (Ichimaru, 1977; Rees et al., 1982; Narayan & Yi, 1994, 1995b; Reynolds et al., 1996). At M87-like mass and accretion rates, radiation at the 230 GHz EHT operational frequency is dominated by synchrotron emission (see, e.g., Yuan & Narayan, 2014). In the synchrotron process, electrons are confined to move in helical orbits about magnetic field lines. This motion sets a characteristic orientation for the electromagnetic fields that are produced and results in a polarization perpendicular to the orientation of the magnetic field lines. In the limit of weak internal Faraday rotation, the observed linear polarization can be used to probe the structure of the local magnetic field.

Accretion flows can be divided into two qualitatively different states according to the properties of their steady-state magnetic fields. In the magnetically arrested disk (MAD) state, the magnetic pressure in the disk near the horizon is large enough to counterbalance the inward ram pressure of the flow (Ichimaru, 1977; Igumenshchev et al., 2003; Narayan et al., 2003). MAD flows are characterized by energetic, quick, violent accretion events and often have higher accretion efficiencies compared to their standard and normal evolution (SANE) counterparts. Since the structure and strength of the magnetic field at the horizon parameterizes the black hole accretion flow state space, linear polarization data may be an efficient discriminator among the underlying models.

Prior work has linked linear polarization fraction and rotation measure in emission from accretion flows to magnetic field structure in simulations of M87 (Broderick & Loeb, 2009; Mościbrodzka et al., 2017). We present a framework for evaluating rotational coherence in the linear polarimetric fields of arbitrary images and use this framework to provide a focused analysis of nearly face-on accretion flow images in a simulated library of M87-like images. We generate a discriminator between the MAD and SANE accretion states and identify trends in the coefficients with respect to black hole spin. The differences in polarization structure

across the image library support a path to inference of the magnetic field structure that may be applicable to EHT polarimetry of M87.

4.3 Decomposition of linear polarization

Images of face-on black hole accretion disks exhibit a ring-like structure that aligns with the symmetry axis of the Kerr spacetime. Although this symmetry persists in the magnitude of polarized intensities, the map of the electric vector position angle (EVPA) depends strongly on the orientation and strength of magnetic fields in the flow. The higher magnetic field strengths present in MADs may lead to increased azimuthal symmetry in EVPA, motivating a symmetry-based decomposition of linearly polarized images to distinguish MAD and SANE states. We describe our method below.

4.3.1 Decomposition definition

For our analysis, we take advantage of the inherent ring-like structure present in black hole images at low inclination by working in polar coordinates (ρ, φ), where the radial distance ρ is measured from the image center and the azimuthal angle φ is measured east of north on the sky. We express the linear polarization on the image in terms of the complex-valued polarization field $P(\rho, \varphi) \equiv Q(\rho, \varphi) + iU(\rho, \varphi)$, where Q and U are the usual Stokes intensities. The corresponding EVPA χ is measured east of north on the sky and can be written in terms of the complex phase of the polarization field $\angle(P)$:

$$\chi = \frac{1}{2} \arctan \frac{U}{Q} = \frac{1}{2} \angle(P). \tag{4.1}$$

We project each image onto a set of basis functions defined in the polarization domain as $P_m(\varphi) \equiv e^{im\varphi}$ to pick out particular modes of azimuthal symmetry. While orthogonal, our basis functions are not complete over the polarimetric image domain because they cannot reproduce radial structure and do not contain absolute polarized flux information without

an accompanying Stokes I image. This is intentional: rather than merely reproduce the polarization map, we wish our decomposition coefficients to be a measure of coherent polarization for particular azimuthal angular dependencies about the image center. Furthermore, because fiducial model images exhibit a sharp ring-like structure, condensing the data along the radial dimension is not expected to lead to significant information loss.

We define the decomposition coefficient β_m to be the scalar product between the basis image and the P image, restricted to an annulus:

$$\begin{aligned}\beta_m &= \frac{1}{I_{\rm ann}} \int_{\rho_{\rm min}}^{\rho_{\rm max}} \int_0^{2\pi} P(\rho,\varphi) P_m^*(\varphi)\ \rho\, {\rm d}\varphi\, {\rm d}\rho \\ &= \frac{1}{I_{\rm ann}} \int_{\rho_{\rm min}}^{\rho_{\rm max}} \int_0^{2\pi} P(\rho,\varphi)\, e^{-im\varphi}\ \rho\, {\rm d}\varphi\, {\rm d}\rho, \qquad (4.2)\end{aligned}$$

$$I_{\rm ann} = \int_{\rho_{\rm min}}^{\rho_{\rm max}} \int_0^{2\pi} I(\rho,\varphi)\ \rho\, {\rm d}\varphi\, {\rm d}\rho\,. \qquad (4.3)$$

We have normalized by $I_{\rm ann}$, the total Stokes I flux in the annulus, and we use $\rho_{\rm min}$ and $\rho_{\rm max}$ to set the radial extent of the annulus.

For each image, we identify the image center and the radial extent of the annulus according to the ring extractor (`rex`) procedure described in detail in Section 9 of EHTC IV. `rex` identifies the ring center as the point that is most equidistant from peak emission along 360 azimuthal slices. The ring width is taken to be the mean full width of half-maximum of the emissivity evaluated along each azimuthal angle from the central point.

For this analysis, we take $\rho_{\rm max} - \rho_{\rm min}$ to be twice the ring width reported by the `rex`-fit ring profile. For GRMHD snapshots blurred to nominal imaging resolutions, this choice ensures that the decomposition coefficients are well-behaved under perturbations to annulus shape and position. Because the nominal image resolution sets the minimum feature size, the area integral will necessarily include at least two distinct resolution elements along any

given radius thereby enabling a meaningful and consistent measure of radial coherence in the polarized image.

Each β_m coefficient is a dimensionless complex number with magnitude corresponding to the amount of coherent power in the m^{th} mode and with phase corresponding to the average pointwise rotation of the image polarization relative to a fiducial EVPA orientation which we define to be vertical along the $\varphi = 0$ image axis. The process can also be thought of as an inner product between the basis functions of Figure 4.2 or as a radially averaged azimuthal Fourier transform of the complex polarization field where the β_m coefficients are Fourier coefficients corresponding to the integral Fourier modes.

If the jet in M87 is aligned with the spin axis of the central black hole as is believed, then our viewing angle to the accretion system should be small (see, e.g., Wang & Zhou, 2009; Mertens et al., 2016; Walker et al., 2018). In such nearly face-on systems, the $m = 2$ mode is of particular interest because of the expected axisymmetric structure in horizon-scale images.

We present a few trivial examples of ring-valued linear polarization fields corresponding to $-4 \leq m \leq 4$ periodic modes with different phases in Figure 4.2. Note that the β_2 coefficient projects out a polarimetric symmetry akin to the E and B modes typically used in studies of polarization in the cosmic microwave background (e.g., Kamionkowski & Kovetz, 2016). In this formalism, the real part of the β_2 coefficient corresponds to E and the imaginary part corresponds to B.

In Figure 4.3, we provide an example of the decomposition applied to images of models that pass EHTC V observational criteria. This example illustrates two strong features of the decomposition. First, blurring increases coherence in polarimetric structure and reduces power in modes with higher $|m|$. Second, the $m = 2$ rotationally symmetric mode can be dominant in both MAD and SANE model snapshots, though in this case, the MAD snapshot has a much larger $m = 2$ coefficient. Furthermore, the complex phase of the β_2 component $\angle(\beta_2)$ encodes the dominant direction of the EVPA spiral. In this MAD snapshot, $\angle(\beta_2) \sim -\pi/2$, corresponding to the EVPA in the bottom right of Figure 4.2,

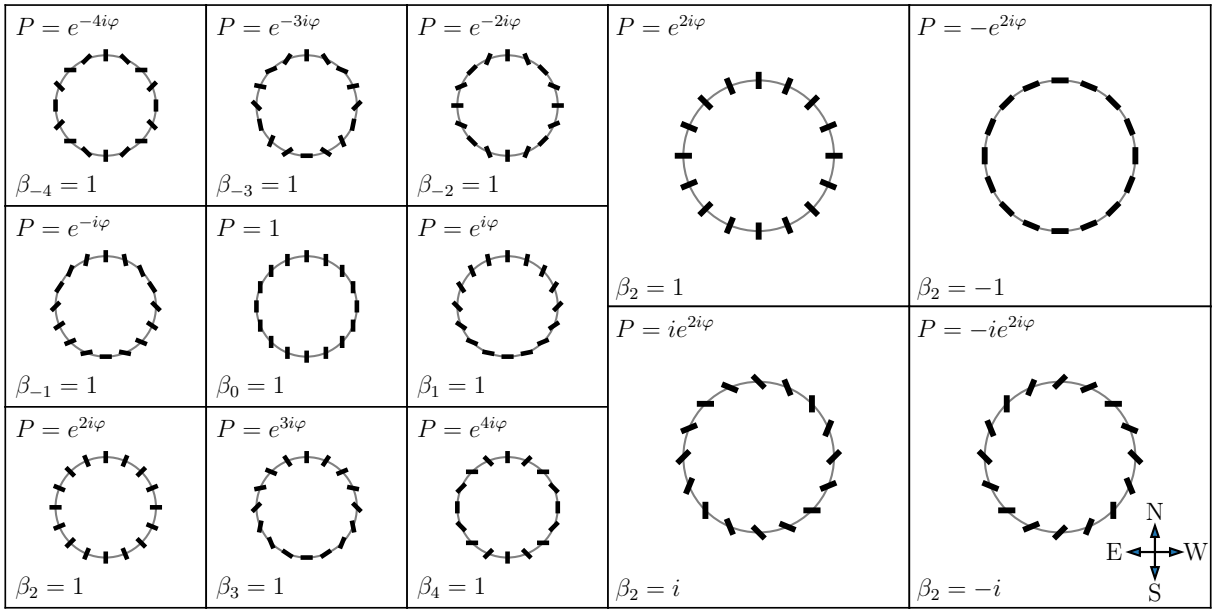


Figure 4.2: Left grid: examples of the electric vector position angle for periodic polarization fields plotted along a ring of unit radius, along with corresponding β_m values for $-4 \leq m \leq 4$. Polarization fields are chosen to produce positive real values of β_m, which correspond to vertical electric vector position angle at the top of the image. Right grid: Same as left, but showing only the rotationally symmetric $m = 2$ mode with four phases in β_2.

while the SANE snapshot has $\angle(\beta_2) \sim 0$, corresponding to a radially directed EVPA.

4.3.2 Interferometric signatures of rotationally symmetric polarization

An interferometer measures visibilities $\widetilde{I}(u, \psi)$ of an image $I(\rho, \varphi)$ according to (see, e.g., Thompson et al., 2017)

$$\widetilde{I}(u, \psi) = \iint I(\rho, \varphi)\, e^{-i2\pi\rho u \cos(\psi - \varphi)}\ \rho\ \mathrm{d}\rho\, \mathrm{d}\varphi, \tag{4.4}$$

where u and ψ give the location of the interferometric baseline when projected onto the image plane perpendicular to the line of sight. Here, u is the magnitude of the baseline vector and ψ is measured east of the positive u_y axis. This definition holds for all Stokes parameters. We adopt the notation $\widetilde{I}$, $\widetilde{Q}$, and $\widetilde{U}$ for the visibilities associated with Stokes

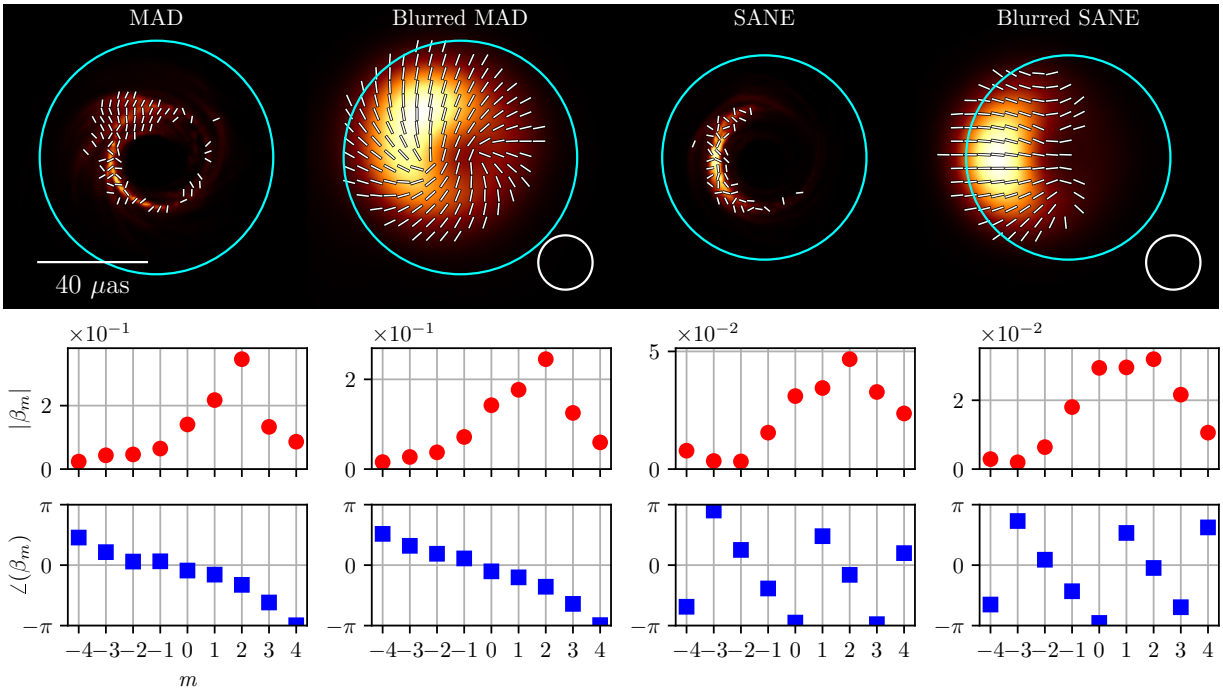


Figure 4.3: Comparison of linear polarimetric decomposition of example MAD and SANE images consistent with observational criteria and approximately equal total flux. The decomposition is applied within the annulus stretching from the blurred `rex`-fit ring diameter to twice its half-width in each direction, centered at the `rex`-fit ring center. Color shows unpolarized Stokes I intensity normalized to unity and ticks show EVPA. Images are shown with and without a 20 μas blurring kernel applied to all Stokes grids. White circles show the blurring kernel; the blue circle shows the outer edge of the `rex` annulus, while the inner edge extends to zero. Both images are of M87-like simulations with a 6.2×10^9 $M_\odot$ central black hole of spin $a_* = 0.94$, viewed at 17° inclination to the black hole spin axis, and with identical models for electron temperature. In figures, EVPA tick marks are shown where both fractional polarization exceeds 1% and Stokes I intensity exceeds 10% of its maximum value. The MAD snapshot is dominated by power in the $m = 2$ mode. Both blurring and integrating over larger scales impose coherence, decreasing power in higher modes.

I, Q, and U respectively. The linear polarization visibilities $\widetilde{P}$ are then $\widetilde{Q} + i\widetilde{U}$ and have interferometric EVPA given by $\frac{1}{2}\angle(\widetilde{P})$.

We now identify the $\widetilde{P}$ signatures corresponding to azimuthal symmetry in P. Consider the simple case of a rotationally symmetric Stokes I image with constant fractional

polarization p. If the polarization evolves azimuthally according to a single mode m, then

$$I(\rho, \varphi) \equiv I(\rho), \tag{4.5}$$

$$P(\rho, \varphi) \equiv pI(\rho)e^{im\varphi}, \tag{4.6}$$

so $P(\rho, \varphi)$ is separable in ρ and φ. The polarized visibilities can then be written

$$\widetilde{P}(u, \psi) = p \int_0^\infty I(\rho) \left[\int_0^{2\pi} e^{im\varphi} e^{-i2\pi\rho u \cos(\psi - \varphi)} \, \mathrm{d}\varphi \right] \rho \, \mathrm{d}\rho \, . \tag{4.7}$$

The integral in φ produces Bessel functions of the first kind J_m and leaves the azimuthal structure intact up to a phase dependence that is determined by m and the sign of J_m

$$\int_0^{2\pi} e^{im\varphi} e^{-i2\pi\rho u \, \cos(\psi - \varphi)} \, \mathrm{d}\varphi = 2\pi i^{-m} J_m(2\pi u\rho) e^{im\psi}. \tag{4.8}$$

Because what remains of Equation 4.7 is an integral in ρ, the angular dependence on $e^{im\psi}$ will be present in the visibility domain and thus dependence on the image angle φ is imprinted on the Fourier domain angle ψ.

The image of a thin ring polarized according to the rotationally symmetric $m = 2$ mode is particularly relevant to our analysis. We can fix the EVPA of the $m = 2$ mode by setting the phase of β_2. One model of a thin polarized ring with diameter d in the image plane is described by Johnson et al. (2020) as

$$I(\rho, \varphi) = \frac{1}{\pi d} \delta\left(\rho - \frac{d}{2}\right), \tag{4.9}$$

$$P(\rho, \varphi) = \beta_2 \, p \frac{1}{\pi d} \delta\left(\rho - \frac{d}{2}\right) \, e^{i2\varphi}. \tag{4.10}$$

The corresponding visibilities are then

$$\widetilde{I}(u,\psi) = J_0(\pi d u), \tag{4.11}$$

$$\widetilde{P}(u,\psi) = -\beta_2 p J_2(\pi d u)\, e^{i2\psi}. \tag{4.12}$$

Evidently, the signature of the $m = 2$ polarization mode is two-fold in $\widetilde{P}$, comprising both the J_2 Bessel function and the $e^{i2\psi}$ azimuthal dependence in the Fourier domain. Each signature is readily identifiable in synthetic data.

Figure 4.4 illustrates these interferometric polarized signatures for three model images: an analytic ring of diameter 40μas whose profile is built from a 1μas rotated boxcar blurred by a 5μas Gaussian, and the MAD and SANE snapshots shown in Figure 4.3. The MAD model shows interferometric signatures similar to those of the $\beta_2 = -i$ right-handed spiral.

Figure 4.4 also shows the overplotted baseline coverage of M87 provided by the 2017 EHT array. Even though the array only provided a sparse sampling of the Fourier plane, we expect that data at the current EHT resolution will still be sensitive to the salient differences in EVPA structure between MAD and SANE models.

4.4 Image library parameter discrimination

By applying the decomposition described in Section 4.3 to ensembles of simulated images and obtaining representative coefficient distributions, we can identify how changing the physical parameters of the black hole accretion system affects the values of the decomposition coefficients. This analysis provides a new domain which might be used to inform parameter extraction efforts. Here, we apply our decomposition to a set of images generated with M87-like parameters. We present the resulting coefficient distributions and then focus on the $m = 2$ coefficient which appears to provide the strongest discriminating power.

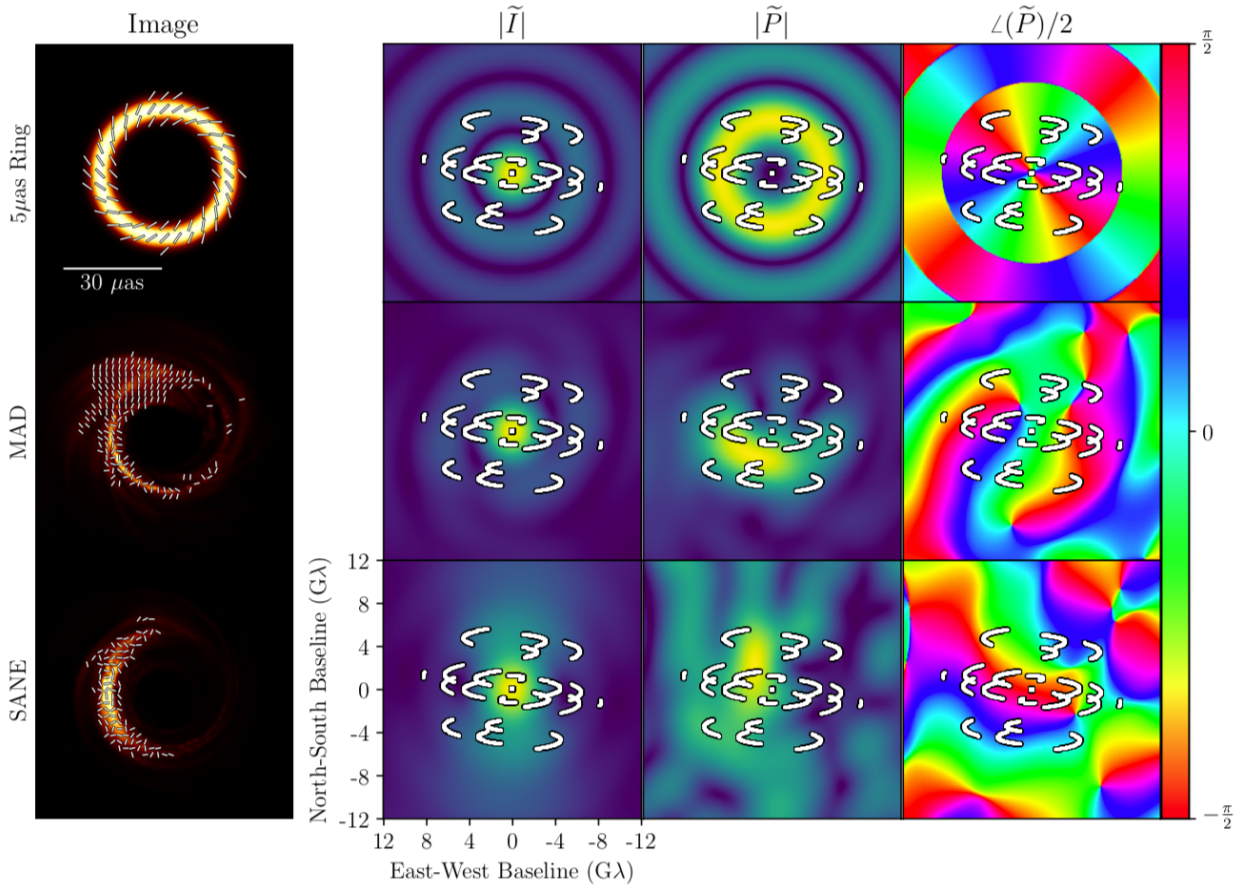


Figure 4.4: Visibility amplitudes $|\widetilde{I}|$ and $|\widetilde{P}|$, and $\frac{1}{2}\angle(\widetilde{P})$ for three example images. The top row presents a 40μas diameter boxcar ring of width 1μas blurred by a 5 μas Gaussian kernel. The bottom two correspond to the GRMHD examples presented in Figure 4.3. EHT 2017 baseline coverage from April 11 is overlaid. Rotational symmetry in the EVPA transfers to the visibility domain, but rotates by 90° on short baselines due to the factor of $-\beta_2 = i$ found in Equation 4.12. Visibility amplitudes are normalized and shown in linear scale.

4.4.1 Image library description

As a part of the analysis effort presented in EHTC V, high-resolution GRMHD simulations of MAD and SANE accretion disks with dimensionless black hole spins $a_* \equiv Jc/GM^2$ ranging between $-0.95 < a_* < 0.99$ were generated. Here, J is the angular momentum of the black hole, and negative values of a_* correspond to anti-parallel black hole and disk angular momentum vectors. Each GRMHD simulation was used to generate a set of $\geq$ 100 polarimetric images evenly spaced in time via general relativistic ray tracing. The ray tracing calculation was performed using the fast light approximation and assumed pure synchrotron

emission and absorption from a thermal electron distribution. As part of the analysis, a set of observational constraints and self-consistency checks was applied to each model. A full description of the library and cuts applied to the models was presented in Chapter 3.

We restrict our analysis to images of black holes with mass $6.2 \times 10^9 M_\odot$ and average compact flux densities $F_\nu(230\,\mathrm{GHz}) = 0.5$ Jy. Each image is generated at an inclination of 17° relative to the angular momentum axis of the black hole in accordance with large-scale estimates of jet inclination (see, e.g., Walker et al., 2018) and in order to produce the characteristic Stokes I brightness asymmetry on the correct side of the image. We take images from both MAD and SANE models with dimensionless spins $\pm 0.94, \pm 0.5$, and 0 with six different electron temperature prescriptions. Our fiducial image set is split evenly among each of the image parameters and includes 3000 each of MAD and SANE images. Of these, 1300 MAD and 600 SANE images pass the observation and consistency checks. The GRMHD simulations we consider were generated with `iharm` (Gammie et al., 2003) and the radiative transfer calculation was performed by `ipole` (Mościbrodzka & Gammie, 2018).

4.4.2 Parameter discrimination results

We apply the decomposition described in Section 4.3 to each image in the library to compute β_m coefficients for $-10 \leq m \leq 10$. Each image is first blurred by a 20 μas Gaussian kernel. The 20 μas size corresponds to slightly less than the nominal beam of the EHT array; EHT imaging algorithms routinely reconstruct images with super-resolution finer than this scale (see EHTC IV; Chael et al., 2016; Kuramochi et al., 2018; Palumbo et al., 2019). We then simultaneously center the image and measure a ring profile using `rex`. Finally, we measure coefficients of the blurred images and examine distributions of the β_m coefficients for MAD and SANE simulations.

Figure 4.5 shows the distribution of the $|m| \leq 4$ coefficients across the selected library images organized by MAD or SANE as well as whether the snapshot belongs to a model that passed the consistency checks. These checks include comparisons to previous observations

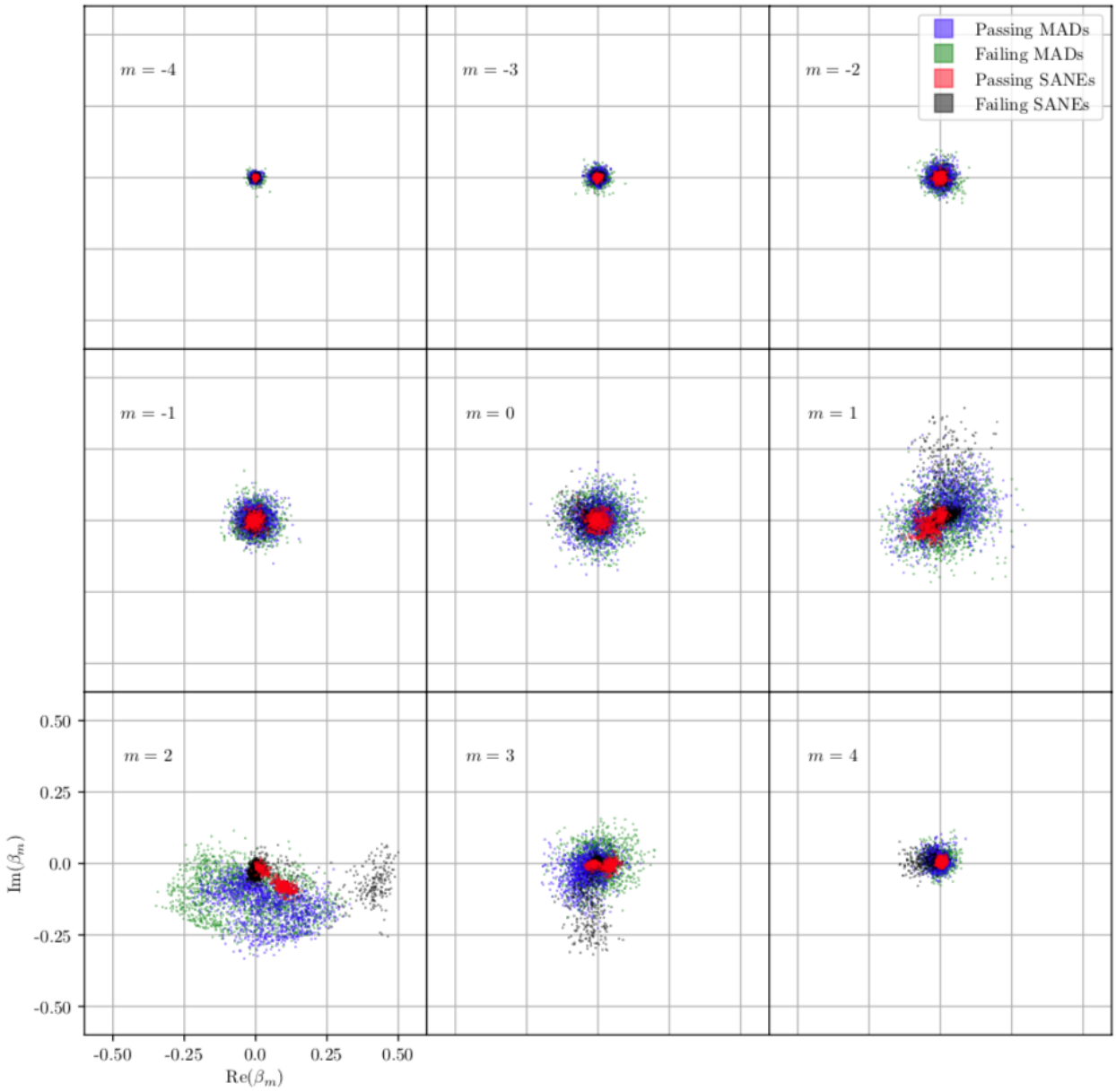


Figure 4.5: Complex β_m coefficients with $-4 \leq m \leq 4$ for the fiducial ray tracing parameters of the GRMHD library after blurring with a 20 μas beam. Coefficients are normalized by the Stokes I annular flux after integrating over a region set by the `rex`-fit radius and width $\rho_{\texttt{rex}} \pm w_{\texttt{rex}}$. Models that are not self-consistent or that are ruled out by prior observational constraints are labeled as failing.

as well as a self-consistency verification that the numerical models do not produce too much radiation to have been accurately simulated by non-radiative GRMHD. Though the $m = 1$ and $m = 3$ modes appear to segregate the MAD and SANE distributions, the largest apparent separation arises in the $m = 2$ mode at the bottom left in the figure.

A small subset (less than 10%) of SANE snapshots exhibit well-ordered polarization fields with even larger β_2 magnitudes than MAD models. These snapshots correspond to prograde spin SANE models in which the ion and electron temperatures are set equal. This choice results in a shift of emission outwards beyond the inner accretion flow and thus EVPA likely traces the magnetic field structure of the disk proper in these models. The original EHT analysis uniformly rejected these models.

The distribution of β_2 magnitudes in Figure 4.6 suggests that discrimination between MAD and SANE is tractable even if the EVPA is arbitrarily rotated. If an image presents a large β_2, it is invariably either a MAD or a failing SANE.

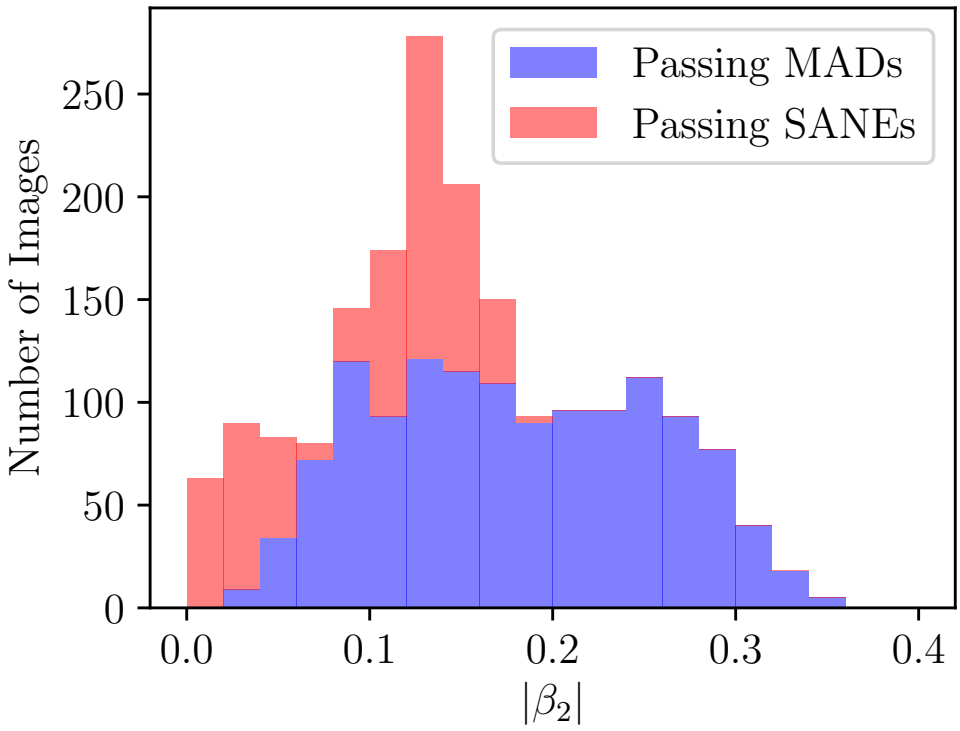


Figure 4.6: Stacked histogram of β_2 coefficient magnitudes shown for MAD and SANE models that are physically consistent with observational criteria. No passing SANE model has $|\beta_2| > 0.20$.

The phase angle of the β_2 coefficient trends with the magnitude of the black hole spin a_*. As can be seen in Figure 4.7, increasing $|a_*|$ correlates with decreasing $\angle(\beta_2)$ and thus a more radially directed EVPA. A coherent external Faraday screen would uniformly rotate

the EVPA map. If the properties of the screen are unknown, constraining spin becomes more challenging.

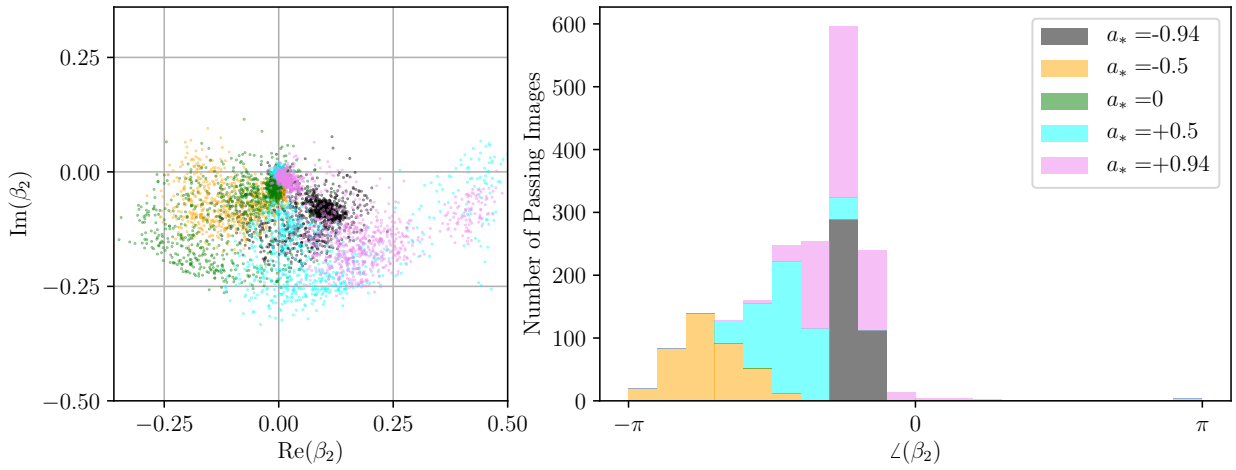


Figure 4.7: Distribution of the complex β_2 coefficient (left) and stacked histogram of the phase of β_2 (right) colored by spin values. Complex coefficients are shown for all models, whereas phases are shown only for passing models. The phase of the β_2 coefficient reflects increasingly radial EVPA at high spins. As shown in the MAD β_2 distribution, magnetic field symmetries in left handed flows correspond to right-handed EVPA maps and corresponding β_2 phases.

The systematically negative phase of the β_2 coefficients is due to an imaging choice. In order to reproduce the ring asymmetry present in the EHT observation, the emitting fluid must be moving clockwise about the hole. This sets a preferred orientation for the black hole angular momentum vector and consequently also the magnetic field. Since EVPA traces magnetic field orientation, this choice preferences right-handed EVPA fields.

4.5 Limitations of the β decomposition model

We have described and validated a decomposition of the linear polarization field into coefficients β_m that quantify ordered symmetries in the image plane. The β_2 coefficient quantitatively identifies rotationally symmetric polarized image structure. The structural sensitivity of the decomposition provides additional constraints beyond measurements of the fractional polarization.

Although it is possible to analytically recover the position of the ring center in simulated images of black holes, we are limited to algorithmic centering procedures for analyses of real data. A polarimetric image of a centered ring with unit flux density given by $P(\varphi) = e^{i(2\varphi+\delta)}$, as in Figure 4.2, will have $|\beta_2| = 1$. In ideal cases, `rex` can correctly identify the image center, enabling accurate computation of the decomposition coefficients. In more general cases, the turbulence in the underlying plasma flow can drastically affect the apparent shape and structure of the observed ring. These discrepancies along with effects due to undersampling in the image domain can result in the `rex` procedure inaccurately identifying the true image center. In orders of the absolute centering error b divided by the ring diameter, b/d, the magnitude of the β_2 coefficient goes like $|\beta_2| \approx 1 - 4\,(b/d)^2$. For rings of finite width w, the correction term gains a factor of $\log(w/d)$. Note, however, that interferometric visibility amplitudes are invariant under centroid shifts, so the structures in Figure 4.4 persist regardless of image centering choices (Thompson et al., 2017).

The EHT Stokes I analysis found full image correlation times to be of order 50 M, or approximately two weeks for M87. In contrast, we find that coherence in the $m = 2$ mode can be significantly longer. Over the full image library, $m = 2$ correlation times ranged from 50 M to nearly 1000 M. The increased coherence results primarily from the blurring and averaging procedures built into the coefficient calculation. Decoherence in the full images is generally driven by hot, isolated fluid features that evolve over a dynamical time. Because these features are localized in azimuth, they do not appear as time dependent features in the $m = 2$ coefficient and thus β_2 retains coherence on longer timescales.

We also found that MAD correlation times are characteristically longer than SANE ones. This is reasonable: because the orientation of the EVPA is connected to the magnetic field structure and because magnetic fields are both stronger and more structured in MAD disks, EVPA orientation should be steadier in MAD disks. Our computation of correlation times is limited by image cadence, and because our cadence is at times of order the calculated coherence time, we have not attempted to provide a more detailed quantification of the

difference between SANEs and MADs by this measure. Figure 4.8 shows the autocorrelation function for both a selected MAD and a selected SANE model. In the case of M87, we expect that images taken at intervals of a year or greater should provide independent realizations of the source and thus improve the statistical accuracy of any probability-based parameter discrimination.

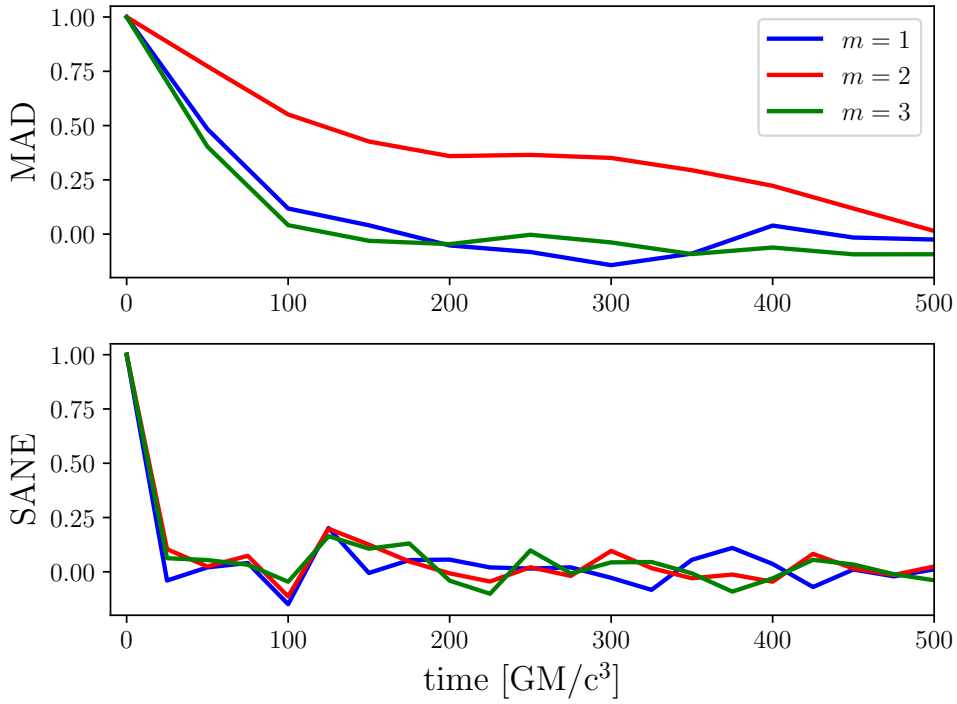


Figure 4.8: Autocorrelation function for several β_m coefficients for an example MAD model and an example SANE model. In these models, MAD correlation times are longer than SANE correlation times. SANE correlation times may not be resolved by the image cadence.

The image library used for this analysis was generated using the fast light approximation in which it is assumed that the time it takes for light to travel through the computational domain is small compared to the timescales on which the fluid properties change. Our analysis is largely insensitive to this choice, especially because the inclination at which we view M87 orients the lines of sight perpendicular to the bulk motion of fluid features.

Aside from the question of fast versus slow light, the values of the decomposition coefficients are geometrically dependent on the orientation at which we view M87. The nearly

face-on view of M87 leads to an inherent symmetry which begets the relative strength of the $m = 2$ mode. Were the same analysis applied to images of black holes viewed edge on at large inclination, the symmetry would be broken. The difference between fast and slow light analyses might also be more pronounced in analyses of black holes viewed closer to edge-on.

Although our analysis was performed entirely within the image domain, images from observations are reconstructions from data products that live in the Fourier plane. Because assumptions about image structure affect the reconstruction procedure (for a review, see EHTC IV), an analysis of polarimetric imaging output based on synthetic data would help ensure that neither (1) imaging choices nor (2) systematic errors due to baseline coverage nor (3) problems in leakage calibration should dominate the signatures of rotational symmetry we have identified.

4.6 Polarimetric observations

4.6.1 Conventions in observations and models

Throughout this paper we use the following definitions and conventions for polarimetric quantities, following the IAU definitions of the Stokes parameters $(\mathcal{I}, \mathcal{Q}, \mathcal{U}, \mathcal{V})$ (Hamaker & Bregman, 1996; Smirnov, 2011). The complex linear polarization field $\mathcal{P}$ is

$$\mathcal{P} = \mathcal{Q} + i\mathcal{U}. \tag{4.13}$$

Then, the electric vector position angle (EVPA) is defined as

$$\text{EVPA} \equiv \frac{1}{2}\arg(\mathcal{P}). \tag{4.14}$$

The EVPA is measured East of North (counter-clockwise from an arrow pointing up) on the sky. Positive $\mathcal{Q}$ is aligned with the North–South direction and negative $\mathcal{Q}$ with the East–West direction. Positive $\mathcal{U}$ is at a +45 deg angle with respect to the positive $\mathcal{Q}$ axis

(in the NE–SW direction). Positive Stokes $\mathcal{V}$ indicates right-handed circular polarization, meaning in our convention that the electric field vector of the incoming electromagnetic wave is rotating counter-clockwise as seen by the observer. In the synchrotron radiation models we consider, a positive value of emitted Stokes $\mathcal{V}$ is associated with an angle θ_B between the wavevector k^μ and magnetic field b^μ as measured in the frame of the emitting plasma in the range $\theta_B \in [0, 0.5\pi]$. Negative $\mathcal{V}$ corresponds to $\theta_B \in [0.5\pi, \pi]$.

The linear and circular polarization fractions at a point in the image are defined as

$$|m| \equiv \frac{|\mathcal{P}|}{\mathcal{I}}, \tag{4.15}$$

$$|v| \equiv \frac{|\mathcal{V}|}{\mathcal{I}}. \tag{4.16}$$

We also define the rotation measure between two wavelengths λ_1 and λ_2

$$\mathrm{RM} \equiv \frac{\mathrm{EVPA}(\lambda_1) - \mathrm{EVPA}(\lambda_2)}{\lambda_1^2 - \lambda_2^2}. \tag{4.17}$$

Unresolved observations measure the *net* (image-integrated) polarization fractions,

$$|m|_{\mathrm{net}} = \frac{\sqrt{\left(\sum_i \mathcal{Q}_i\right)^2 + \left(\sum_i \mathcal{U}_i\right)^2}}{\sum_i \mathcal{I}_i}, \tag{4.18}$$

$$v_{\mathrm{net}} = \frac{\sum_i \mathcal{V}_i}{\sum_i \mathcal{I}_i}, \tag{4.19}$$

where the sums are over all pixels i in the resolved image. In addition to the signed circular polarization fraction v_{net}, we also frequently consider the absolute value $|v_{\mathrm{net}}|$, since circular polarization measurements of the M87 core at 230 GHz do not constrain its sign (Goddi et al., 2021).

In describing the *resolved* linear polarization in EHT images, we define the image-average

linear polarization fraction, weighted by the total intensity of each image pixel, as

$$\langle |m| \rangle = \frac{\sum_i \sqrt{\mathcal{Q}_i^2 + \mathcal{U}_i^2}}{\sum_i \mathcal{I}_i}. \tag{4.20}$$

Note that $\langle |m| \rangle$ depends on the imaging resolution (beam size), while $|m|_{\rm net}$ is the usual unresolved linear polarization fraction and does not depend on resolution.

4.6.2 Unresolved polarization and rotation measure measurements towards M87's core from ALMA

As part of the EHT 2017 observation campaign, we obtained ALMA array measurements of the unresolved, net ∼230 GHz polarimetric properties of M87's core and jet on April 5, 6, 10, and 11 (hereafter these observations are referred to as ALMA–only observations). These results, along with details on the observations and data calibration, are presented in Goddi et al. (2021); we summarize them here in Table 4.1. From the ALMA-only data, the net linear polarization fraction (Equation 4.18) of the core is $|m|_{\rm net} \simeq 2.7\%$. The data also provide an upper limit on the net circular polarization fraction (Equation 4.19) of the core of $|v|_{\rm net} \lesssim 0.3\%$, with a magnitude and sign that vary over the four observed epochs. Goddi et al. (2021) also measured an EVPA that varies with wavelength across the ALMA band; the slope of EVPA with wavelength is consistent with EVPA $\propto \lambda^2$, as expected for Faraday rotation. The inferred rotation measure (Equation 4.17) is time variable and changes sign between April 5 and 11, with a maximum magnitude $|{\rm RM}| \simeq 1.5 \times 10^5\,{\rm rad\,m^{-2}}$.

The ALMA-only measurements include extended ∼arcsecond scale structures that are entirely resolved out of the EHT maps of M87's core region. As a result, the total 230 GHz flux density of M87 measured by ALMA alone is a factor of $\simeq 2$ larger than that captured by the resolved EHT images (see also EHTC IV).

Day	F (Jy)	$\lvert m\rvert_{\rm net}$ (%)	$\lvert v\rvert_{\rm net}$ (%)	RM ($10^5\,{\rm rad\,m^{-2}}$)
April 5	1.28 ± 0.13	$2.42 \pm .03$	≤ 0.2	(0.6 ± 0.3)
April 6	1.31 ± 0.13	$2.16 \pm .03$	≤ 0.3	(1.5 ± 0.3)
April 10	1.33 ± 0.13	$2.73 \pm .03$	≤ 0.3	(-0.2 ± 0.2)
April 11	1.34 ± 0.13	$2.71 \pm .03$	≤ 0.4	(-0.4 ± 0.2)

Table 4.1: ALMA–only measurements of M87's unresolved polarization properties at $\nu = 221$ GHz (Goddi et al., 2021).

4.6.3 Spatially resolved linear polarization of M87's core in EHT 2017 data

The resolved polarimetric images of the M87 core reported in EHTC VII display robust features between different image reconstruction algorithms and across four days of observations (April 5, 6, 10 and 11 of 2017). At $20\,\mu$as resolution, the images consistently show a region of highest linear polarized intensity in the south-west portion of the ring, with a fractional linear polarization $|m| \lesssim 30$ % at its maximum. The image-average linear polarization fraction takes on values $5.7\% \leq \langle |m| \rangle \leq 10.7\%$ across the different observation days and image reconstruction techniques. The range of the image-integrated net polarization fraction is $1.0\% \leq |m|_{\rm net} \leq 3.7\%$ (see EHTC VII, Table 2). Because polarized emission outside the EHT field-of-view but inside the ALMA-only core is unconstrained, we adopt the EHT $|m|_{\rm net}$ range when evaluating models rather than the ALMA-only values.

On all four observing days, the EHT images reveal a characteristic azimuthal pattern of the EVPA angle around the emission ring. To quantify this pattern across the image, we use the PWP method to decompose $\mathcal{P} = \mathcal{Q} + i\mathcal{U}$ into azimuthal modes with complex coefficients β_m as described earlier in this chapter and in Palumbo et al. 2020. For a polarization field in the image plane given in polar coordinates (ρ, φ), the β_m coefficients are

$$\beta_m = \frac{1}{I_{\rm ann}} \int_{\rho_{\rm min}}^{\rho_{\rm max}} \int_0^{2\pi} \mathcal{P}(\rho, \varphi)\, e^{-im\varphi}\, \rho\, {\rm d}\varphi\, {\rm d}\rho, \tag{4.21}$$

where $I_{\rm ann}$ is the Stokes I flux density contained inside the annulus set by the limiting radii $\rho_{\rm min}$ and $\rho_{\rm max}$. We take $\rho_{\rm min} = 0$ and $\rho_{\rm max}$ to be large enough to include the entire EHT image.

Within the library of polarized images from GRMHD simulations produced for EHTC V, PWP found that the $m = 2$ coefficient, β_2, was the most discriminating in identifying the underlying magnetized accretion model. The phase of β_2 maps well onto the qualitative behavior expected of polarization maps with idealized magnetic field configurations. In our convention, radial EVPA patterns have positive real β_2 ($\angle\beta_2 = 0\,{\rm deg}$), azimuthal EVPA patterns have negative real β_2 ($\angle\beta_2 = 180\,{\rm deg}$), and left (right) handed spiral patterns have positive (negative) pure imaginary β_2 ($\angle\beta_2 = 90\,{\rm deg}$ and $-90\,{\rm deg}$, respectively). The measured range of the complex β_2 coefficient across the different image reconstruction methods and observing days reported in EHTC VII, Table 2 is: $0.04 \leq |\beta_2| \leq 0.07$ for the amplitude, and $-163\,{\rm deg} \leq \arg[\beta_2] \leq -129\,{\rm deg}$ for the phase.

Trends in β_2 metric computed across the GRMHD image library (Section 4.8) can be obtained in the visibility domain using only E- and B- mode measurements taken on EHT 2017 baselines, as long as the visibilities are accurately phase calibrated (see Appendix A of EHTC VIII). Since accurate phase calibration of EHT data is non-trivial and requires fully modeling the polarized source structure, we use image domain comparisons to the reconstructions presented in EHTC VII for the constraints applied in the rest of this paper.

As in total intensity, both the unresolved and resolved polarimetric properties of the 230 GHz M87 image changed over the week between April 5 and April 11. Notably, the integrated EVPA in the EHT image changes by $\approx 30 - 40\,{\rm deg}$ (while the ALMA-only EVPA changes by $\lesssim 10\,{\rm deg}$). We will not interpret this variability further in this work, although Section 4.11 provides a broad overview of expectations for time variability in viable models. The observational ranges of the key parameters we use in comparing theoretical models to data in Section 4.9—namely $|m|_{\rm net}$, $|v|_{\rm net}$, $\langle|m|\rangle$, and β_2 amplitude and phase—are summarized in Table 4.2.

Table 4.2: Parameter ranges for the quantities used in scoring theoretical models in this paper.

Parameter	Min	Max
$\|m\|_{\rm net}$	1.0%	3.7 %
$\|v\|_{\rm net}$	0	0.6 %
$\langle\|m\|\rangle$	5.7 %	10.7 %
$\|\beta_2\|$	0.04	0.07
$\angle\beta_2$	$-163\,{\rm deg}$	$-129\,{\rm deg}$

Note – The ranges for $|m|_{\rm net}$, $\langle|m|\rangle$, and β_2 were taken from EHTC VII Table 2. These ranges represent the minimum -1σ error bound and maximum $+1\sigma$ error bound across five different image reconstruction methods, and incorporate both statistical uncertainty in the polarimetric image reconstruction and systematic uncertainty in the assumptions made by different reconstruction algorithms. The upper limit on $|v|_{\rm net}$ was taken as $\simeq 2\times$ the value from Goddi et al. (2021).

4.6.4 External and internal Faraday rotation

Faraday rotation in a uniform plasma with rotation measure RM rotates the EVPA away from its intrinsic value EVPA_0 according to Equation 4.17. The change in EVPA from its intrinsic value at 230 GHz ($\lambda \simeq 1.3\,{\rm mm}$) is:

$$\Delta\mathrm{EVPA} \simeq 9.7 \left(\frac{\mathrm{RM}}{10^5\,{\rm rad\,m^{-2}}}\right) {\rm deg}\,. \tag{4.22}$$

Polarized light rays passing through a uniform medium are subject to the same RM. The net source polarization angle is then coherently rotated away from its intrinsic value without any depolarization. This scenario of *external* Faraday rotation has been used to infer the mass accretion rate for sources where an RM is measured or constrained (e.g., Bower et al. 2003; Marrone et al. 2006, 2007; Kuo et al. 2014), by assuming that the observed radiation passes through the bulk of the accretion flow. Since relativistic electrons suppress the Faraday rotation coefficient as $\propto 1/T_e^2$ (e.g., Jones & Hardee, 1979), these models assume the RM is produced outside the emission region at the radius where $\Theta_e = kT_e/m_ec^2 = 1$, usually $r \sim 100\,r_g$ (where $r_g = GM/c^2$ is the gravitational radius).

However, in accreting systems like M87, it is unclear whether this external Faraday

rotation model is a good approximation. As we estimate below, one zone emission models of M87 predict substantial RM within the emission region itself at radii $r \lesssim 5\, r_g$. At its low viewing inclination, it is also unclear whether the observed polarized radiation emitted near the horizon passes through any cool, high density infalling gas. Therefore, *internal* Faraday rotation, which can depolarize the emission as well as rotate the EVPA (Burn, 1966), is also an important effect to consider.

The observed $\simeq 15\%$ linear polarization of the ring at the EHT scale of $\sim 20\,\mu$as is much lower than the intrinsic synchrotron polarization fraction $\gtrsim 70\%$ expected locally. This could result from synchrotron self-absorption of the emitted radiation, but one-zone estimates and theoretical models (e.g., EHTC V, and references therein) suggest that the system at 230 GHz is most likely optically thin. It is more likely that the observed depolarization of the resolved emission could be the result of polarization structure that is scrambled at resolutions finer than the EHT beam. Turbulent magnetic fields and Faraday rotation internal to the emission region could produce this scrambling. In Section 4.8.3 we show that turbulence in GRMHD models alone is insufficient to produce this level of depolarization. Significant internal Faraday rotation of polarization vectors on different rays by different amounts can produce a sufficiently scrambled image that is depolarized when spatially averaged over a telescope resolution element/beam (e.g., Burn, 1966; Agol, 2000; Quataert & Gruzinov, 2000; Beckert & Falcke, 2002; Ruszkowski & Begelman, 2002; Ballantyne et al., 2007).

From the simultaneous ALMA-only M87 observations, the RM implied by changes in the EVPA across the ALMA band is $|\mathrm{RM}| \lesssim 1.5 \times 10^5\,\mathrm{rad\,m^{-2}}$. These values are consistent with, but much more constraining than, the range determined from past SMA observations (Kuo et al. 2014). The ALMA-only EVPA difference varies by order unity in magnitude and sign over the observing campaign and includes a large flux contribution from extended emission not captured by EHT 2017 imaging (EHTC IV). Using a two-component model, Goddi et al. (2021) shows that the rotation measure toward the core emission in the EHT field-of-view is uncertain and can exceed the rotation measure computed from all of the

ALMA-only data including emission that is resolved out by the EHT, with values as large as $|RM| \lesssim 10^6\,\mathrm{rad\,m^{-2}}$. Because of this uncertainty, we do not use the observed rotation measure as an observational constraint in our analysis. We account for uncertainty related to the observed time variability by using reconstructed polarized EHT images from both April 5 and 11 to define the acceptable ranges Table 4.2 of the observational parameters used to score theoretical models in Section 4.9.

4.7 Estimates

In this section, we take a first look at the importance of internal Faraday rotation and magnetic field structure in determining the characteristics of the 230 GHz EHT image. In §4.7.1, we obtain order-of-magnitude estimates of the plasma properties in M87 by interpreting the observed depolarization as entirely due to the effect of internal Faraday rotation on small scales. In Section 4.7.2 we explore the effects of different idealized magnetic field configurations on the observed polarization pattern from plasma orbiting a black hole in the absence of Faraday effects.

4.7.1 Parameter estimates from one-zone models

Based on a one-zone isothermal sphere model, EHTC V derived order-of-magnitude estimates of the plasma number density n_e and magnetic field strength B in the emitting region around M87. Constrained by the Stokes $\mathcal{I}$ image brightness, size, and total flux density, they found

$$n_e \simeq 2.9 \times 10^4\,\mathrm{cm^{-3}}, \tag{4.23}$$

$$B \simeq 4.9\,\mathrm{G}. \tag{4.24}$$

In this model, the emission radius was assumed to be $r \simeq 5r_g$, and the electron temperature was assumed to be $T_\mathrm{e} = 6.25 \times 10^{10}$ K, based on the observed brightness temperature of

the EHT image. This temperature corresponds to $\Theta_e = kT_e/m_e c^2 = 10.5$, so the emitting electrons have moderately relativistic mean Lorentz factors $\bar{\gamma} \approx \Theta_e/3 = 3.5$. The angle between the magnetic field and line-of-sight is set at $\theta = \pi/3$. This model ignores several physical effects that are included in more sophisticated models and simulations and which are necessary to fully describe the emission from M87. The plasma is considered to be at rest and so there is no Doppler (de)boosting of the emitted intensity from relativistic flow velocities. Redshift from the gravitational potential of the black hole is also not included.

Given $n_{\rm e}$, B and $T_{\rm e}$, we can estimate the strength of the Faraday rotation effect at 230 GHz quantified by the optical depth to Faraday rotation τ_{ρ_V}:

$$\tau_{\rho_V} = r \times \rho_V \simeq 5.2 \left(\frac{r}{5 r_g} \right), \tag{4.25}$$

where ρ_V is the Faraday rotation coefficient (e.g., Jones & Hardee, 1979). For emission entirely behind an external Faraday screen, τ_{ρ_V} is related to the rotation measure RM via $\tau_{\rho_V} = 2{\rm RM}\lambda^2$, which follows from the radiative transfer equations for spherical Stokes parameters in the absence of other effects (see, e.g., Appendix A of Mościbrodzka et al., 2017) and the fact that $\rho_V \propto \lambda^2$.

Our estimated τ_{ρ_V} indicates that Faraday rotation internal to the emission region is an important effect and could thus explain the depolarization observed in M87. Faraday effects are even more important for the case of polarized light emitted by relativistic electrons that travel through a dense, colder accretion flow (e.g., Mościbrodzka et al., 2017; Ricarte et al., 2020). In addition, for the same parameters, Faraday conversion of linear to circular polarization may also be important ($\tau_{\rho_Q} \simeq 0.5$), while self-absorption is weak ($\tau_I \simeq 0.05$). Requiring an internal Faraday optical depth $\tau_{\rho_V} > 2\pi$ (large enough to produce significant depolarization) provides an additional constraint on one-zone models independent of those used in EHTC V, which fixed the electron temperature at an assumed value. Assuming $\tau_{\rho_V} > 2\pi$ allows us to break the degeneracy between magnetic field strength, electron temperature, and plasma number density.

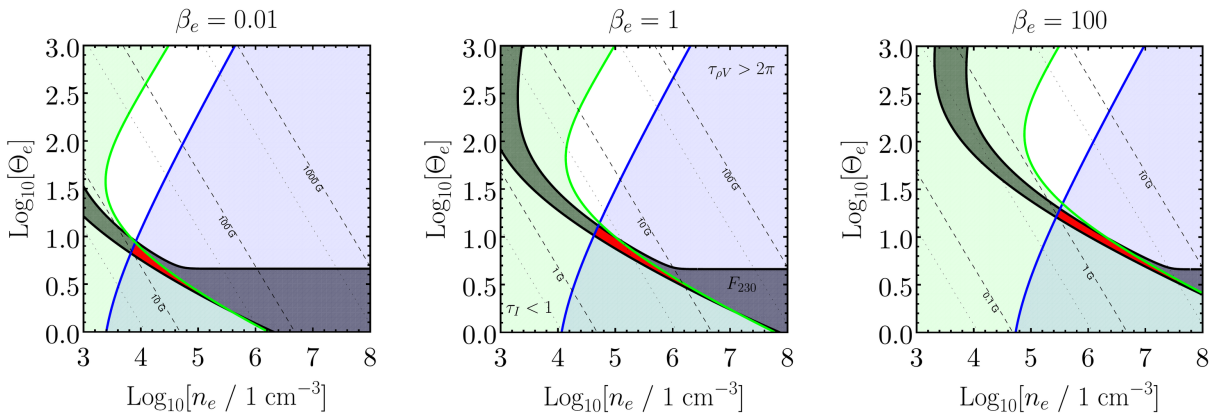


Figure 4.9: Allowed parameter space in number density and dimensionless electron temperature (n_e, Θ_e) (red region) for the simplistic one-zone model described in subsection 4.7.1 for three constant values of $\beta_e = 8\pi n_e m_e c^2 \Theta_e / B^2$. We require that the optical depth $\tau_I < 1$ (green region), the Faraday optical depth $\tau_{\rho_V} > 2\pi$ (blue region), and the total flux density $0.2 < F_\nu < 1.2$ Jy (black region). Contours of constant magnetic field strength are denoted by labeled dashed lines.

Hence, we consider the same model as in EHTC V at several different values of $\beta_e = 8\pi n_e k T_e / B^2$, constrained by the requirement that the Faraday optical depth $\tau_{\rho_V} > 2\pi$. To be consistent with the 230 GHz EHT data, we also require that the observed image have a total flux F_ν between 0.2 and 1.2 Jy, and that the model has a maximum total intensity optical depth $\tau_I < 1$. Figure 4.9 shows what constraints these requirements put on the electron number density n_e and the dimensionless electron temperature Θ_e at three different values of β_e. For values of $0.01 < \beta_e < 100$, the electron temperature in this simple model is constrained to lie in a mildly relativistic regime $2 \lesssim \Theta_e \lesssim 20$ ($10^{10} < T_e < 1.2 \times 10^{11}$ K), and the magnetic field strength is $1 \lesssim B \lesssim 30$ G. The number density of the emitting electrons depends more sensitively on the assumed value of β_e, taking on values between $10^4\,\mathrm{cm}^{-3}$ and $10^7\,\mathrm{cm}^{-3}$.

The one-zone model estimates suggest that both the total intensity and polarized emission can be produced in a mildly relativistic plasma in a magnetic field of relatively low strength $B \lesssim 30$ G. For higher values of B, the electron temperature would be too small to explain the observed maximum brightness temperature ($\simeq 10^{10}$ K) in the M87 EHT image (EHTC IV). Very high values of B are independently disfavored by the small degree of circular polarization

$|v|_{\rm net} \lesssim 1\%$ seen in M87. For $B \simeq 100$ G, the ratio of the Stokes $\mathcal{V}$ emissivity to the Stokes $\mathcal{I}$ emissivity $j_V/j_I \simeq 1\%$. For $B \simeq 10^3$ G, $j_V/j_I \simeq 10\%$, for all temperatures $> 10^{10}$ K. We also note that magnetic fields of $B \gtrsim 5$ G are sufficient to produce jet powers of $P_{\rm jet} \gtrsim 10^{42}\,{\rm erg\,s^{-1}}$ (e.g., EHTC V) via the Blandford & Znajek (1977) process.

4.7.2 EVPA pattern and field geometry

To demonstrate how the intrinsic magnetic field structure in the emission region influences the observed polarization pattern, we present the polarization configurations from three idealized magnetic field geometries around a black hole—a purely toroidal field, a purely radial field, and a purely vertical field—as seen by a distant observer. In Figure 4.10 we show polarimetric images from these simple field configurations computed with two methods: a numerical model of an optically thin emission region around the black hole (top row of Figure 4.10), and an analytic treatment of the parallel transport of the polarization vector that is originally perpendicular to the magnetic field (Narayan et al. 2021, middle row of Figure 4.10). We show the polarization maps from both methods for the three idealized magnetic field configurations viewed at an inclination angle of $i = 163$ deg. Both the analytical and numerical calculations assume a zero spin black hole (Schwarzschild metric), though we have found that the main features of these polarization patterns are insensitive to spin.

In the top row of Figure 4.10 we show the result of numerical calculations performed with the general relativistic ray tracing code `grtrans` (Dexter & Agol, 2009; Dexter, 2016) of polarized emission from an optically and Faraday thin compact emission region, or "hotspot", in Keplerian orbit around a black hole in the equatorial plane. The hotspot has a radial extent of $3r_g$ and moves in an imposed and idealized magnetic field geometry in a circular orbit at a radius of $8r_g$ (following Gravity Collaboration et al., 2018, 2020b). We construct a phenomenological model of a torus of emitting, rotating plasma by studying the time-averaged polarized emission images from one revolution of this hotspot around the black hole. We have verified that a semi-analytic implementation (Broderick & Loeb, 2006) of a

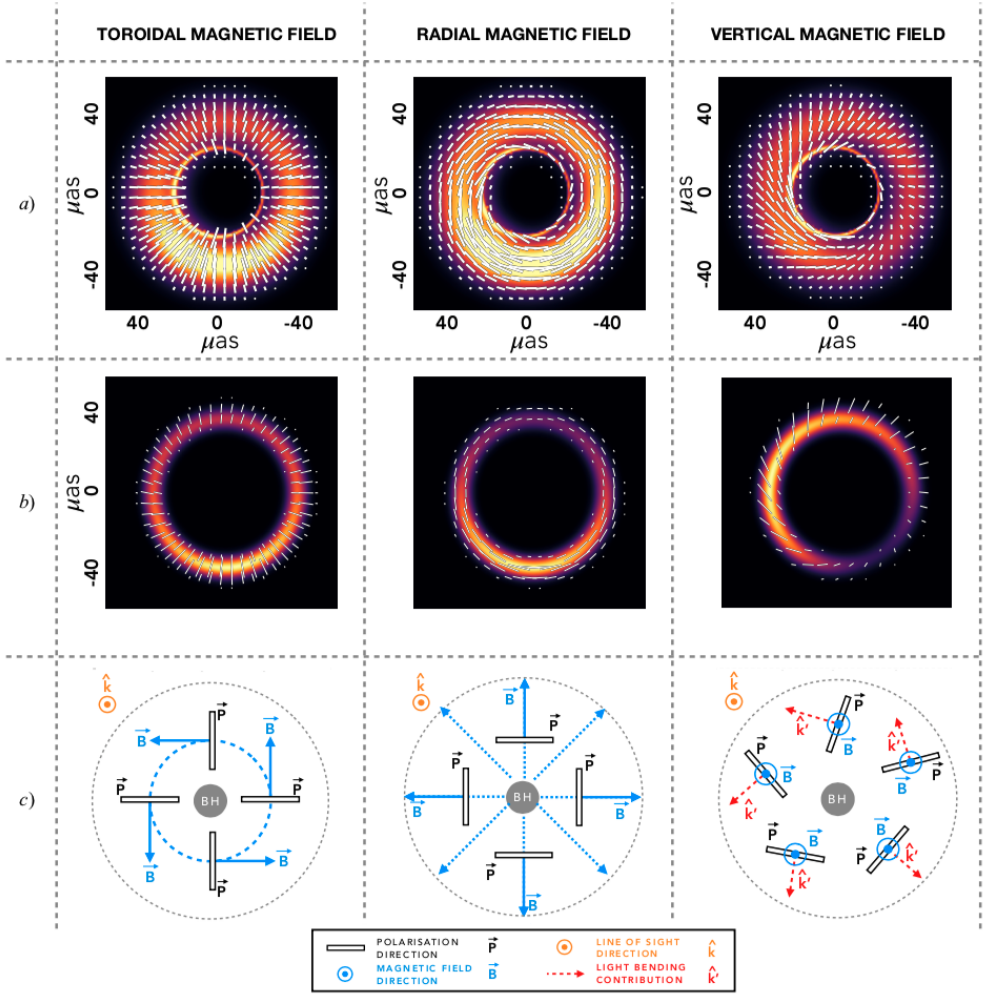


Figure 4.10: *a*) Numerical calculations of the polarization configuration generated by an orbiting emission region in the shape of a torus at $8r_g$ in three imposed magnetic field geometries and viewed at $i = 163\,\mathrm{deg}$ (with material orbiting clockwise on the sky). The orbital angular momentum vector is pointing away from the observer and to the east (to the left). Total intensity is shown in the background with higher brightness temperature regions shown as lighter in color. In the foreground, the observed EVPA direction is shown with white ticks, with the tick length proportional to the polarized flux. *b*) Analytic calculations of the polarization configuration from a thin ring of magnetized fluid at $8r_g$ inclined by 163 deg to the observer in the same magnetic field geometries as in *a*). While the distribution of emitting material is different in the two models, both the sense of asymmetry in the brightness distributions and the polarization patterns match those from the numerical calculations. *c*) Schematic cartoons showing the emitting frame wave-vector k, magnetic field direction $\vec{B}$, and polarization vector $\vec{P} = \hat{k} \times \vec{B}$ for each case. In the bottom right panel, $\hat{k}'$ denotes the approximate light bending contribution to the wave-vector.

hot accretion flow model (Yuan et al., 2003) also produces consistent polarization patterns when using the same field geometry.

In the second row of Figure 4.10, we compare these numerical results to results from an analytic calculation of the observed polarization pattern generated by the emission of polarized light on a thin ring of radius $8r_g$ in the equatorial plane. In this model (Narayan et al., 2021), the polarization vectors are emitted perpendicular to the imposed magnetic field geometry in the fluid rest frame; they are transformed on their way to the observer using an approximate, but analytic treatment of the effects of light bending, parallel transport, and Doppler beaming. This calculation includes radial inflow as well as rotation in the velocity field; the models shown use purely toroidal motion (clockwise on the sky) with the same idealized magnetic field geometries as in the numerical case. The models match the asymmetric brightness distributions and polarization patterns of the numerical calculations. In particular, both models produce a consistent helical EVPA pattern in the case of a vertical magnetic field.

The linear polarization direction $\bar{P}$ of synchrotron radiation in the emitted frame is perpendicular to the wave-vector $\hat{k}$ and the magnetic field vector $\bar{B}$. We define the toroidal magnetic field as consisting only of magnetic field components in the azimuthal direction, while the poloidal magnetic field consists of the remainder, including both radial and vertical components. In a purely toroidal field case, the EVPA pattern shows a radial EVPA pattern (left column in Figure 4.10). Purely radial magnetic fields (middle column) give a complementary result; the polarization has a toroidal configuration, similar to a 90 deg rotation of the linear polarization ticks from the toroidal case.

In a vertical magnetic field (right column in Figure 4.10), we might expect that $\bar{P}$ should be vertical (North-South) everywhere since a vertical $\bar{B}$ is tilted East-West for this viewing geometry. We might also expect that $\bar{P} \simeq 0$ when the black hole is viewed close to face on, since $\hat{k}||\bar{B}$. Instead, the linearly polarized emission from a purely vertical field shows a twisting pattern that wraps around the black hole. This twist results from light bending in

the emitting region near the black hole, which provides a radial contribution $\hat{k}'$ to the emitted wave-vector $\hat{k}$ that initially points away from the black hole (see the schematic cartoon in the bottom right panel of Figure 4.10). As a result, close to the black hole, the total wave-vector $\hat{k}_{\rm emit} = \hat{k} + \hat{k}'$ and the magnetic field $\bar{B}$ are no longer parallel, and the polarization is non-zero. In addition, the relativistic motion of the emitting material (aberration) breaks the North-South symmetry from a pure vertical field and gives the twisting pattern a handedness corresponding to the orbital direction. The EVPA patterns in these images do not show a strong dependence on the black hole spin.

In a rotating flow, weak magnetic fields are sheared into a predominantly toroidal configuration (e.g., Hirose et al., 2004). In the absence of other effects (e.g., external Faraday rotation), the observed azimuthal EVPA pattern suggests the presence of dynamically important magnetic fields in the emission region, which can retain a significant poloidal component in the presence of rotation.

In the next sections, we compare numerical simulations of the accretion flow and jet-launching region in M87 with different field configurations to the EHT2017 data to better constrain the magnetic field structure.

4.8 M87 model images from GRMHD simulations

The low resolved fractional linear polarization observed by the EHT contradicts the results from an idealized magnetic field structure with no disorder. For typical parameters of the 230 GHz emission region, Faraday rotation and conversion are expected to be important. Magnetic field structure, plasma dynamics and turbulence, and radiative transfer effects including Faraday rotation can be realized in images from three-dimensional general relativistic magnetohydrodynamic (3D GRMHD) simulations of magnetized accretion flows. We use 3D GRMHD simulations (described in Section 4.8.1) in combination with polarized general relativistic radiative transfer (GRRT) models (described in Section 4.8.2) to model polarized

images of M87. In Section 4.8.3, we describe trends of the key observables ($|m|_{\rm net}$, $|v|_{\rm net}$, $\langle|m|\rangle$, and β_2) in our GRMHD polarimetric image library.

4.8.1 GRMHD model description

The simulation library generated for the analysis of the EHT 2017 total intensity data in EHTC V consists of a set of 3D GRMHD simulations that were post-processed to generate simulated black hole images via GRRT. For simulations using black holes with non-zero angular momentum, we only considered accretion flows in which the angular momentum of the flow and the hole were aligned (parallel or anti-parallel). Since the equations of non-radiating[1] GRMHD are scale invariant, each fluid simulation was thus fully parameterized by two values describing the angular momentum of the black hole and the relative importance of the magnetic flux near the horizon of the accretion system.

The black hole angular momentum J is expressed in terms of the dimensionless black hole spin parameter $a_* \equiv Jc/GM^2$. In this paper, we consider simulations run with the `iharm` code (Gammie et al., 2003; Noble et al., 2006) with $a_* = -0.94$, -0.5, 0, 0.5, and 0.94, where positive (negative) spin implies alignment (anti-alignment) between the accretion disk and the black hole angular momentum. Several studies of "tilted" disks have been conducted (e.g. Fragile et al., 2007; McKinney et al., 2013; Morales Teixeira et al., 2014; Liska et al., 2018; White et al., 2019; Chatterjee et al., 2020). As there does not yet exist a full library of tilted disk simulations spanning a range of spins and the computational cost associated with generating such a library is high, we limit our analysis to the aligned and anti-aligned simulations considered in EHTC V.

The strength of the magnetic flux near the horizon qualitatively divides accretion flow solutions into two categories: the Magnetically Arrested Disk (MAD) state (e.g., Bisnovatyi-Kogan & Ruzmaikin, 1974; Igumenshchev et al., 2003; Narayan et al., 2003) in which the

[1]We assume that M87 is well described by models in which radiative cooling is negligible so that it does not affect the dynamics of the plasma and model images of M87 can be generated using post-processing radiative transfer models.

magnetic flux near the horizon saturates and significantly affects the dynamics of the flow, and the contrasting Standard and Normal Evolution (SANE) state (e.g., Gammie et al., 2003; De Villiers & Hawley, 2003; Narayan et al., 2012). The relative importance of magnetic flux in a simulation is quantitatively described by the dimensionless quantity:

$$\phi \equiv \Phi_{\rm BH} \left(\dot{M} r_g^2 c \right)^{-1/2}, \tag{4.26}$$

where $\Phi_{\rm BH}$ is the magnitude of the magnetic flux crossing one hemisphere of the event horizon (see Tchekhovskoy et al., 2011; Porth et al., 2019) and $\dot{M}$ is the mass accretion rate through the event horizon. The flux saturates at values of $\phi \gtrsim 50$, at which point the flow becomes MAD. The SANE simulations we consider have lower values of $\phi \approx 5$.[2] Accreted material supplied at large scales could in principle supply any value of net vertical flux. Here, we do not explore cases with small or zero net vertical flux $\Phi_{\rm BH} \lesssim 1$. We also do not consider values in the relatively narrow intermediate range $5 \lesssim \Phi_{\rm BH} \lesssim 50$.

The SANE simulations we consider were run on a grid resolution of $288 \times 128 \times 128$, with a fluid adiabatic index $\gamma = 4/3$, and an outer simulation domain of $r_{\rm out} = 50\, r_g$. The MAD simulations had a grid resolution of $384 \times 192 \times 192$, an adiabatic index $\gamma = 13/9$, and an outer simulation domain of $r_{\rm out} = 10^3\, r_g$. The simulations were carried out in the FMKS set of coordinates described in §2.7, which concentrate resolution near the midplane and away from the pole at small radii.

4.8.2 Ray-traced polarimetric images from GRMHD simulations

Unlike the equations of GRMHD, the equations of radiative transfer are not scale invariant, so we must introduce a length scale and a mass/density scale when we ray-trace images from the numerical fluid data. The length (and time) scale is set by the mass of the black

[2]Note that the MAD threshold $\phi \gtrsim 50$ is given in Gaussian units where $[\Phi] = \mathrm{G\,cm^2}$. If the field strength is given in the Lorentz–Heaviside units typically used in simulations ($B_{\rm LH} = B_{\rm G}/\sqrt{4\pi}$), the MAD threshold on the dimensionless flux $\phi \sim 15$.

hole, assumed to be $M_{\rm BH} = 6.2 \times 10^9 M_\odot$ in accordance with the value used to generate the EHTC V simulation library. For our models, we also adopt the $D = 16.9$ Mpc distance to M87 used in EHTC V. The density scale of the accreting plasma (equal to the scale of the magnetic pressure) is chosen so that on average the simulated images reproduce the observed 230 GHz compact flux density, $F_\nu \simeq 0.5$ Jy.

Images were generated from the set of simulations over several values of the polar inclination angle i that were chosen to be broadly consistent with observational estimates of the inclination angle of the M87 jet (e.g., Walker et al., 2018). The position angle on the sky can be changed after image generation by rotating both the image and the Stokes $\mathcal{Q}$ and $\mathcal{U}$ components appropriately. Each image has a 320×320 pixel resolution over a $160\,\mu$as field of view, where each pixel contains full Stokes $\mathcal{I}, \mathcal{Q}, \mathcal{U}, \mathcal{V}$ intensities.

In GRMHD simulations, we make the approximation that the plasma is thermal, i.e., that the electrons and ions are described by a Maxwell–Jüttner distribution function (Jüttner, 1911). However, the plasma around M87 and in other hot accretion flows is most likely collisionless, with electrons and protons that are unable to equilibrate their temperatures (e.g. Shapiro et al., 1976; Ichimaru, 1977). We mimic collisionless plasma properties in producing images from the GRMHD simulations by allowing the electron temperature $T_{\rm e}$ to deviate from the proton temperature $T_{\rm i}$. The simulations used in this work only track the total internal energy density $u_{\rm gas}$, not the distinct electron and ion temperatures. We set $T_{\rm e}$ after running the simulation according to local plasma parameters following the parameterization introduced by Mościbrodzka et al. (2016, see also EHTC V). The ratio between the ion and electron temperatures R is determined by the local plasma $\beta = p_{\rm gas}/p_{\rm mag}$, where $p_{\rm gas} = (\gamma - 1)u_{\rm gas}$, and $p_{\rm mag} = B^2/8\pi$. The temperature ratio is then taken to be

$$R = \frac{T_i}{T_e} = R_{\rm high}\frac{\beta^2}{1+\beta^2} + R_{\rm low}\frac{1}{1+\beta^2}, \tag{4.27}$$

where $R_{\rm high}$ ($R_{\rm low}$) are the free parameters of the model and give the approximately constant

temperature ratio at high (low) β. This approach allows us to associate the electron heating with magnetic properties of plasma.

In calculating the electron temperature, we further assume that the plasma is composed purely of ionized Hydrogen, that the ions are nonrelativistic with an adiabatic index $\gamma_p = 5/3$, and that the electrons are relativistic with $\gamma_e = 4/3$.

We note that this procedure is not entirely self-consistent, since the γ of the combined electron-ion fluid will change depending on the relative pressure contributions of electrons and protons while we assume it is fixed throughout the simulation domain. See Sądowski et al. (2017) for an alternative, self-consistent approach.

In this paper, we consider a library of 72000 simulated images composed of sets of 200 realizations of the same accretion system described by a fixed set of heating/observation parameters. Each set of 200 images is drawn from output files spaced by $25 - 50\,r_g/c$ from the set of ten GRMHD simulations spanning five spin values in both MAD and SANE field configurations. The inclination angle for each image is set to one of either $i = 12, 17, 22\,\mathrm{deg}$ (retrograde models, $a_* < 0$) or $i = 158, 163, 168\,\mathrm{deg}$ (prograde models, $a_* \geq 0$), according to the parity that is required to orient the brightest portion of the ring in the southern part of the image while ensuring the position angle of the approaching jet is consistent with large scale observations.

We use electron heating parameters $R_{\rm low} = 1, 10$ and $R_{\rm high} = 1, 10, 20, 40, 80$, or 160 in Equation 4.27. EHTC V only considered models with $R_{\rm low} = 1$. Larger values of $R_{\rm low}$ correspond to lower electron-to-proton temperature ratios in the low β regions (e.g., the jet funnel). This choice is physically motivated for M87, where radiative cooling of the electrons may keep $T_e < T_i$ even in magnetized regions where electron heating is efficient (e.g., Mościbrodzka et al., 2011; Ryan et al., 2018; Chael et al., 2019). Lower electron temperatures in $R_{\rm low} = 10$ models increase the Faraday rotation depth and can result in increased depolarization in parts of the image.

GRMHD simulations produce a highly magnetized jet funnel above the black hole's poles

away from the accretion disk. In the funnel, where the plasma magnetization parameter $\sigma \gg 1$, our numerical methods typically fail to accurately evolve the plasma internal energy. In the image library, we cut off all emission in regions where $\sigma > 1$ to ensure that we limit the emitting region to plasma whose internal energy is safely evolved without numerical artifacts (as in EHTC V). We tested the importance of a $\sigma > 1$ electron population by generating a supplementary set of images from all models with a cut at $\sigma = 10$ and found that it did not change the overall distribution of the derived metrics we use for model scoring in Section 4.9.

Each set of 200 model images with the same parameters in the image library requires a unique density scaling factor that is determined by matching the average flux density from the model to the observed compact flux density of M87 measured by the EHT. Hence, the mass accretion rates, radiative efficiencies, and jet powers will differ between two models even if they are derived from the same underlying simulation (e.g., if $R_{\rm high}$, $R_{\rm low}$, or i is changed). The additional models discussed in Section 4.11, which explore the effects of different σ cutoff values and the inclusion non-thermal electrons, also require unique mass scaling factors.

All of the polarimetric images that we analyze in this paper were generated using the `ipole` code (Mościbrodzka & Gammie, 2018). A comparison of contemporary GRRT codes can be found in Gold et al. (2020). A preliminary comparison of contemporary polarized transfer codes show consistency at the fraction of 1% in all Stokes parameters. All calculated images in this work ignore light travel time delays through the emission region (they adopt the so-called "fast light" approach) and are calculated at a single frequency $\nu = 230$ GHz, neglecting the finite observing bandwidth of the EHT. We confirm that neither of these effects are important for models of interest for M87.

4.8.3 Sample GRMHD model images and polarization maps

We show images and polarization maps for a subset of library models in Figure 4.11. In general, since the horizon scale magnetic fields in MAD models are strong enough not to

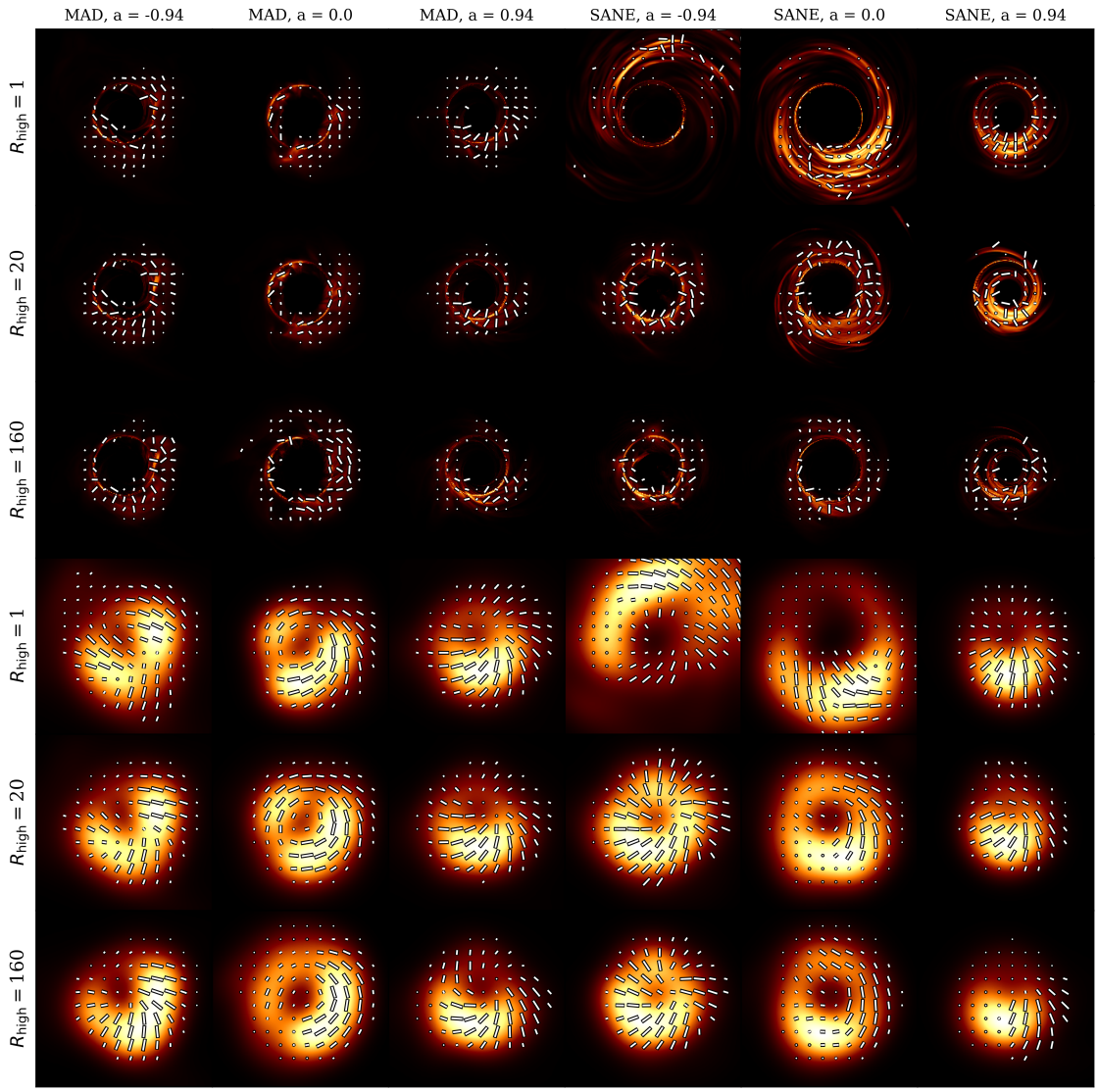


Figure 4.11: Sample snapshot false color images and polarization maps for a subset of the models in the EHT M87 simulation image library at their native resolution (top three rows) and blurred with a 20 μas circular Gaussian beam (bottom three rows). The inclination angle for all images is either 17 deg (for negative a_* models) or 163 (for positive a_* model) deg, with the black hole spin vector pointing to the left and away from the observer. The tick length is proportional to the polarized flux, saturated at 0.5 of the maximum value in each panel. Here models with $R_{\rm low} = 1$ are shown. In general, the EVPA pattern is predominantly azimuthal for MAD models (e.g., MAD $a_* = 0$ $R_{\rm high} = 1$) and radial for SANE models (e.g. SANE $a_* = 0.94$ $R_{\rm high} = 1$), although the SANE $a = 0$ models in particular are exceptions to this trend. All models show scrambling in the polarization structure on small scales from internal Faraday rotation, with more pronounced scrambling in models with cooler electrons (larger $R_{\rm high}$ parameter).

be advected with the accretion flow, they are more likely to have a significant poloidal component and produce azimuthal EVPA patterns (Figure 4.10). In contrast, SANE models tend to show more radial EVPA patterns. Some MAD $a_* = 0.94$ and SANE $a_* = 0$ images are notable exceptions to this trend. These trends are also apparent in the distributions of the β_2 phase across the full image library we consider later in Figure 4.15.

The GRMHD models at their native resolution include notable disorder in the EVPA structure, resulting from both magnetic turbulence and Faraday rotation. Models with larger $R_{\rm high}$ have lower electron temperatures and higher Faraday rotation depths, resulting in the most disordered polarization maps. Many of the EVPA patterns seen in the images that have been blurred with a 20 μas Gaussian kernel resemble those from the idealized magnetic field models in Figure 4.10, indicating that the net EVPA pattern after blurring may trace the intrinsic magnetic field structure.

In Figure 4.12 we show a sample polarization map at full resolution compared to the same map blurred with circular Gaussian kernels of 10 μas and 20 μas FWHM. From tests with synthetic data, blurring (convolving with a circular Gaussian kernel) provides a reasonable approximation to image reconstruction from the EHT data at a comparable resolution (EHTC VII). The resolved average fractional polarization in the blurred images $\langle |m| \rangle$ traces the degree of order in the intrinsic polarization map. In the blurred images, disordered polarized structure on small scales produces beam depolarization. The degree of depolarization decreases with increasing spatial resolution (decreasing beam size).

The bottom row of Figure 4.12 shows the same unblurred and blurred polarization maps, but calculated without the effect of Faraday rotation ($\rho_V = 0$). Those images show more coherent EVPA structure, with much larger $|m|_{\rm net}$ and, particularly when blurred, much larger $\langle |m| \rangle$. Evidently, for this particular model, the depolarization visible in the corresponding upper panels is due to Faraday rotation internal to the emission region. In addition, the net EVPA pattern shifts by a significant amount. The change in β_2 by $\simeq 80\,$deg would correspond to an apparent RM of $\simeq -4 \times 10^5\,{\rm rad\,m^{-2}}$. Our GRRT calculations include all

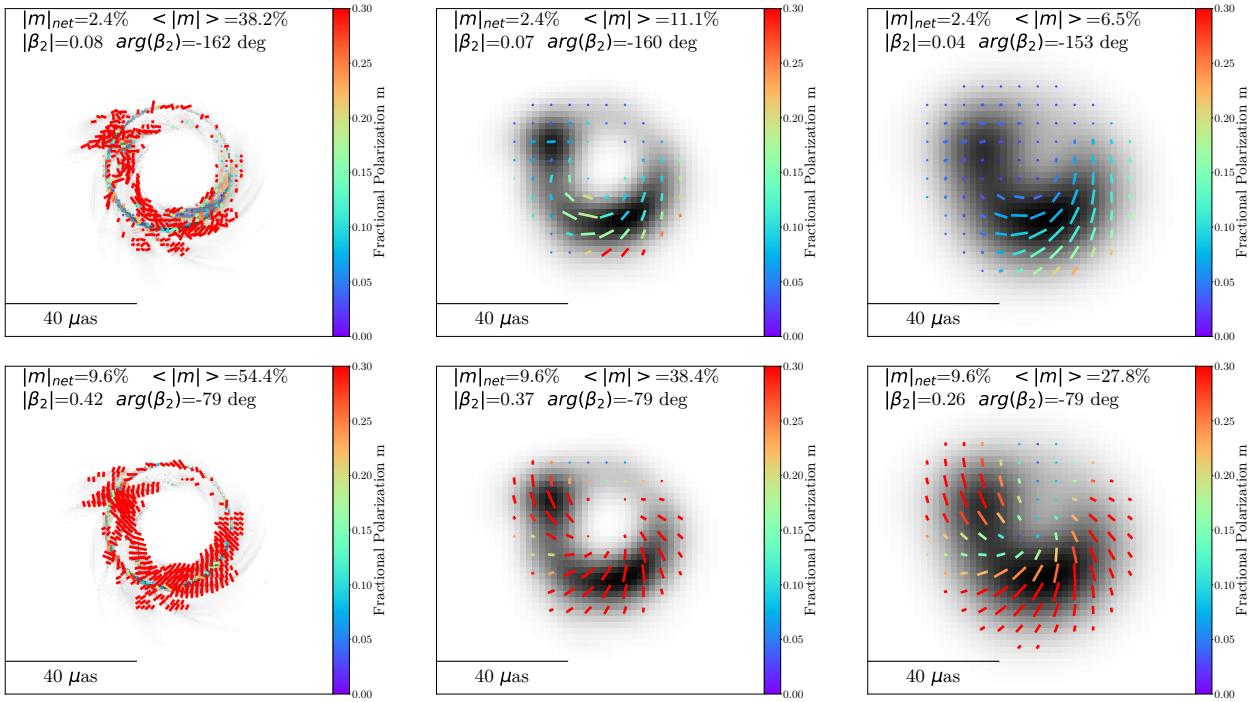


Figure 4.12: Left: a sample polarization map from the image library at original resolution, taken from the MAD $a_* = 0.5$ ($R_{\rm low} = 10$, $R_{\rm high} = 80$) model. Middle and right: the same map but convolved with a 10 μas and 20 μas FWHM circular Gaussian beam, respectively. The position angle of the black hole spin in all frames is PA $= +90$ deg and the inclination angle is $i = 158$ deg, meaning that the black hole spin points left and away from the observer. The lower panels show the same model but calculated with $\rho_V = 0$ (no Faraday rotation). When Faraday rotation is excluded, the EVPA pattern is more coherent, resulting in much larger values of $|m|_{\rm net}$ and $\langle|m|\rangle$. There is also a net rotation of the EVPA pattern between the two cases, by $\simeq 80$ deg in the phase of β_2.

Faraday rotation occurring inside the GRMHD simulation domain ($r_{\rm out} = 50 - 100\,r_g$), both external and internal to the 230 GHz emission region. The observables considered here, for the low viewing inclination of M87, do not depend strongly on that outer radius, as long as it is at $r \gtrsim 40\,r_g$. We cannot rule out the presence of additional Faraday rotating material at larger radii $\gtrsim 100\,r_g$, and its effects are not included in our models.

4.8.4 GRMHD model theory metrics

We compute the polarimetric observables ($|m|_{\rm net}, |v|_{\rm net}, \langle|m|\rangle, \beta_2$) described in Section 4.6.3 from model images blurred with a circular Gaussian kernel with a FWHM of 20 μas in order to compare them to the ranges measured from EHT and ALMA-only data. Both $\langle|m|\rangle$ and β_2 depend on the resolution and hence the size of the Gaussian blurring kernel. Notice also that the value of β_2 also depends on the choice of the image center. When comparing with the range inferred from the EHT image reconstructions (which have been centered by aligning them to the centered, fiducial total intensity images), we do not shift the library images before computing β_m coefficients. Figure 4.13 (right panel) shows the resolved average polarization fraction $\langle|m|\rangle$ as a function of their image averaged Faraday rotation depth, $\langle\tau_{\rho V}\rangle$. At small $\langle\tau_{\rho V}\rangle$, the average polarization fraction is $\langle|m|\rangle \simeq 20 - 50\%$. Intrinsic disorder in the magnetic field structure due to turbulence is generally insufficient to produce the low observed image average polarization fraction in EHT 2017 M87 data ($5.7\% \leq \langle|m|\rangle \leq 10.7\%$). This is especially evident for the SANE models with prograde black hole spin, which have the highest resolved polarization fractions. At large $\langle\tau_{\rho V}\rangle$, strong scrambling from internal Faraday rotation typically results in small predicted polarization fractions of $< 5\%$ at the scale of the EHT beam.

The clear exception to this trend are some SANE retrograde models ($a_* = -0.9375$ for large $R_{\rm high}$), which show $\langle|m|\rangle \simeq 10 - 20\%$ despite their large $\tau_{\rho V} \gtrsim 10^3$. In these models, most of the observed polarized flux originates in the forward jet, while most of the computed Faraday depth is accumulated near the midplane. Photons that travel from the forward jet

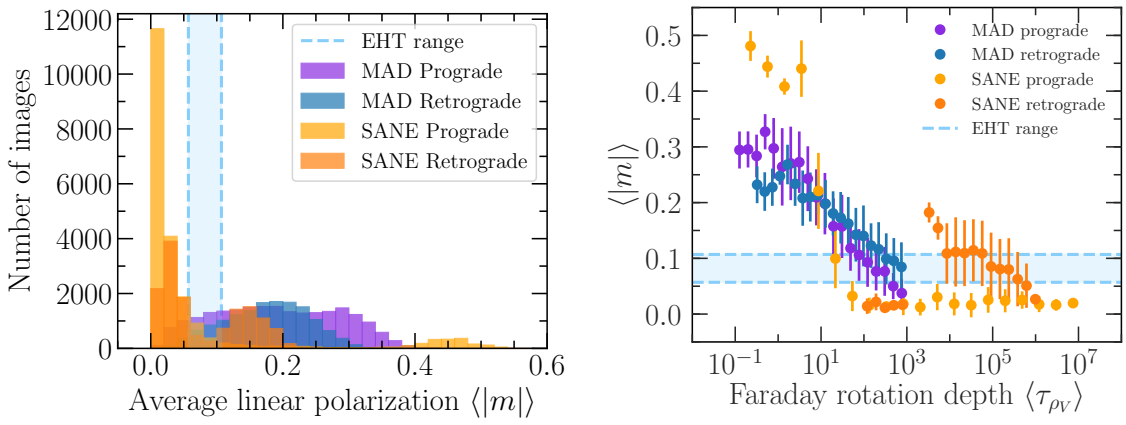


Figure 4.13: Left: distribution of image-averaged fractional polarization $\langle|m|\rangle$ over the M87 library images blurred with a 20 μas beam. The measured range from reconstructed polarimetric images of M87 is shown in dashed lines. Right: $\langle|m|\rangle$ as a function of the intensity-weighted Faraday depth across each image for library images blurred with the same 20 μas circular Gaussian beam. The Faraday depth is calculated as the intensity-weighted sum of $|\rho_V|$ integrated along each ray and increases monotonically with increasing $R_{\rm high}$ for fixed values of the other parameters. A large Faraday depth corresponds to scrambling of the polarization map, which decreases the coherence length of the EVPA (Jiménez-Rosales & Dexter, 2018). Increased scrambling results in stronger depolarization at the scale of the EHT beam and lower values of $\langle|m|\rangle$.

to the observer do not encounter the large Faraday depth. For similar reasons, the inferred RM can be much lower than implied by their large values of integrated τ_{ρ_V}.

Distributions of all observables are shown in Figure 4.13 ($\langle|m|\rangle$, left panel), Figure 4.14 ($|m|_{\rm net}$ and $|v|_{\rm net}$), and Figure 4.15 ($|\beta_2|$ and $\angle\beta_2$). SANE models tend to have a lower integrated polarization fraction and larger circular polarization fraction than M87 at 230 GHz. In many cases this is a result of very large Faraday rotation internal to the emission region. MAD models tend to have larger net linear polarization fraction than observed in M87. The resolved average fractional polarization produces similar trends. Most SANE models with prograde spin are too scrambled and most MAD models are too ordered compared to the reconstructed polarization maps of M87. Full distributions for all models, including their dependence on $R_{\rm high}$, $R_{\rm low}$, and a_*, are discussed in § 4.10.

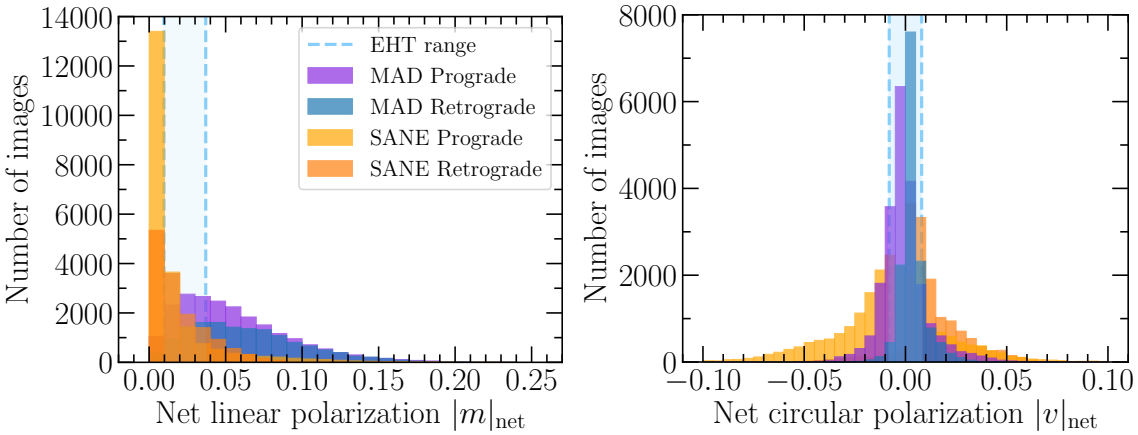


Figure 4.14: Distributions of image-integrated net linear (left) and circular (right) polarization fractions for all EHT M87 library images. The dashed lines show the allowed range inferred from EHT image reconstructions (for $|m|_{\rm net}$) and ALMA–only data (for $|v|_{\rm net}$).

4.9 Model evaluation

4.9.1 Model constraints from polarimetry

To evaluate whether a given GRMHD model is consistent with the EHT observations reported in EHTC VII, we require images from the model to satisfy constraints on the four parameters derived from the reconstructed EHT images and ALMA-only measurements presented in Table 4.2 and summarized again here.

1. The image-integrated net linear polarization $|m|_{\rm net}$ is in the measured range from the EHT image reconstructions: $1\% \leq |m|_{\rm net} \leq 3.7\%$.

2. The image-integrated net circular polarization $|v|_{\rm net}$ satisfies an upper limit from ALMA-only measurements reported in Goddi et al. (2021): $|v|_{\rm net} \leq 0.6\%$.

3. The image-averaged linear polarization $\langle|m|\rangle$ is in the measured range from the EHT image reconstructions at $20\,\mu$as scale resolution: $5.7\% \leq |m|_{\rm net} \leq 10.7\%$.

4. The amplitude and phase of the complex β_2 coefficient quantifying coherent azimuthal structure fall in the measured range: $0.04 \leq |\beta_2| \leq 0.07$ and $-163^\circ \leq \arg[\beta_2] \leq -129^\circ$.

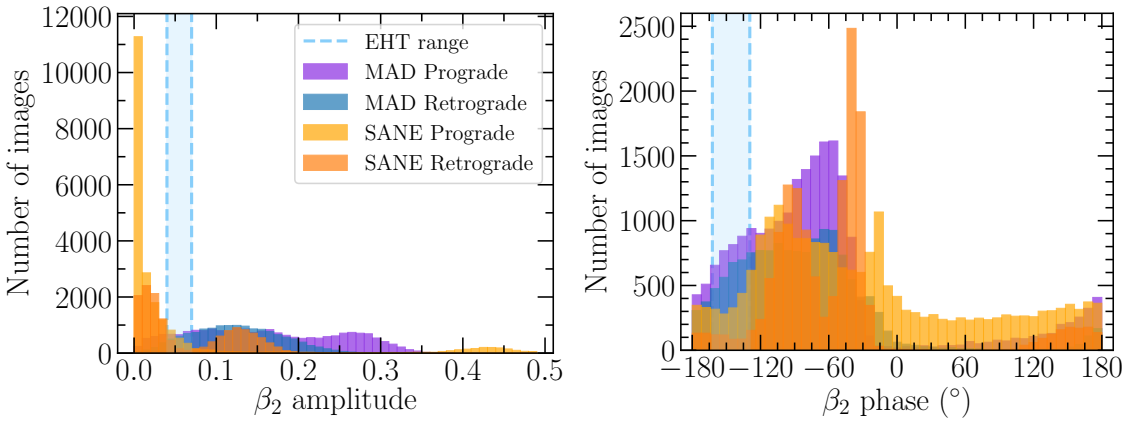


Figure 4.15: Distributions of β_2 amplitude (left) and phase (right) for EHT M87 library images blurred with a 20 μas beam. The measured ranges from reconstructed images of M87 are shown as dashed lines.

We use 72000 library images (from Section 4.8) with 200 time snapshots per model at three inclination angles, six values of $R_{\rm high} = 1, 10, 20, 40, 80, 160$, two values of $R_{\rm low} = 1, 10$, five values of $a_* = -0.9375, -0.5, 0, +0.5, +0.9375$, and realized with both MAD and SANE magnetic field configurations.

In comparing models to observables, the β_2 metric is the most constraining. Only 790 snapshot images out of 72000 considered fall in the range of those reconstructed in both β_2 amplitude and phase, compared to 11,526 snapshots for both $|m|_{\rm net}$ and $|v|_{\rm net}$ and 7,727 for the resolved image-average linear polarization fraction $\langle|m|\rangle$.

Below we explore two quantitative methods for scoring models, either by requiring that at least one single snapshot image from a model simultaneously passes all constraints (*simultaneous* scoring, Section 4.9.2) or that each observational constraint is satisfied by at least one snapshot image from a given model (*joint scoring*, Section 4.9.3).

4.9.2 Simultaneous snapshot model scoring

In the simultaneous scoring procedure, we rule out models where none of the 600 snapshot images (200 time samples at 3 inclination angles) can simultaneously satisfy the constraints on all of the polarimetric observables. Only 73/72000 snapshot images across 15/120 models

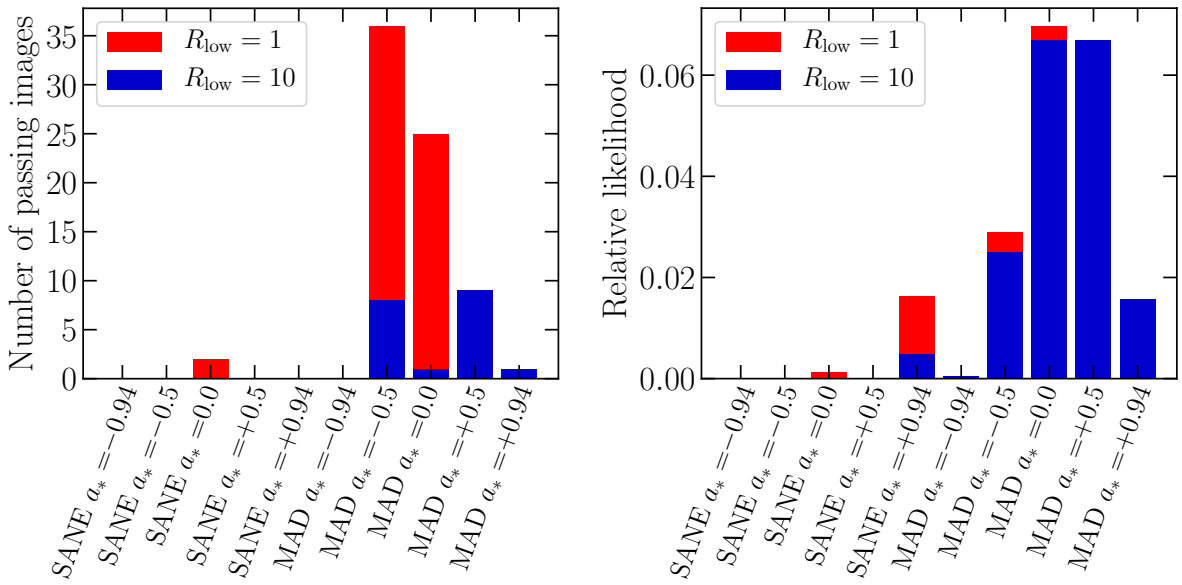


Figure 4.16: Results of the simultaneous (left) and joint (right) scoring methods for comparing GRMHD models to M87 observables. The simultaneous scoring method shows the total number of viable images for each image library model after summing over $R_{\rm high}$. Out of a total of 73 passing images, only 2 are from a SANE model. All passing images are from models with $R_{\rm high} \geq 20$. The right panel shows the joint likelihood of each library model after summing over $R_{\rm high}$. In this method, $R_{\rm low} = 10$ MAD models are preferred and SANE $a_* = +0.94$, $R_{\rm high} = 10$ models are also allowed.

simultaneously pass all of the constraints. Of those, all but 2 viable snapshot images come from a MAD model. The only models with more than 5 passing images are MAD $a_* = 0$ $R_{\rm low} = 1$ $R_{\rm high} = 160$ and MAD $a_* = -0.5$ $R_{\rm low} = 1$ $R_{\rm high} = 80$, 160.

Visually, images that pass the simultaneous scoring procedure show good qualitative agreement with the primary features of the EHT image in Figure 4.1. In contrast, the snapshots from the ruled-out models tend to be too polarized, too depolarized, or too radial in their EVPA pattern. There is variation within individual snapshots in each of the models, but the systematic differences over the 5 observables we consider enables us to effectively constrain the models overall. The left panel of Figure 4.16 shows the total number of images that pass simultaneous scoring as a function of model, summing over the six $R_{\rm high}$ values.

4.9.3 Joint distribution model scoring

In the alternative joint scoring procedure, we use the measured distributions of the data metrics to ask whether the observed value of each metric for M87 is consistent with being drawn from the distribution seen in the GRMHD simulations. To do this, we measure χ^2 values for the five metrics $x_j \in \{|m|_{\rm net}, |v|_{\rm net}, \langle|m|\rangle, |\beta_2|, \angle\beta_2\}$ for all snapshots k from a given model as

$$\chi^2_{j,k} = \frac{(x_{j,k} - \bar{x}_j)^2}{\sigma_j^2}, \tag{4.28}$$

where $x_{j,k}$ are the values of a scoring metric x_j for each of the 600 snapshots k from a given model, $\bar{x}_j$ is the mean of those values for the model, and σ_j is taken as one half of the *observed* data range from Table 4.2. Note that the scoring results of this method do not depend on the choice of σ_j. We then calculate an analogous $\chi^2_{j,\rm data}$ value for the midpoint of the measured range from Table 4.2. A likelihood value $\mathcal{L}_j$ of the data being drawn from the model distribution is defined as the fraction of images with $\chi^2_{j,k} > \chi^2_{j,\rm data}$. The joint likelihood of each model is the product $\mathcal{L} = \Pi_j \mathcal{L}_j$ of those for the 5 metrics x_j.

To produce a non-zero likelihood $\mathcal{L}$ in this method, at least one snapshot from a model must lie further from its mean than the data value does. That can be a different snapshot for each metric, which makes this method more lenient than the simultaneous scoring method. We also note that snapshots are allowed to have the *wrong sign* of the difference with their mean, due to the definition of χ^2 and our use of the mean of the model snapshots themselves. In practice, this makes little difference in the results.

In this method, we consider models viable whose joint likelihood is $> 1\%$ of the maximum found from any model. The right panel of Figure 4.16 shows the resulting joint likelihoods summed over $R_{\rm high}$.

4.9.4 Comparison of scoring results

The results of both scoring procedures are summarized in Figure 4.16, summed over $R_{\rm high}$. Both scoring methods prefer MAD models to SANE models, with most of the passing models coming from the MAD $a_* = 0$ and $a_* = \pm 0.5$ simulations.

The main difference between the two scoring procedures is that joint scoring prefers $R_{\rm low} = 10$ models, while $R_{\rm low} = 1$ is preferred by simultaneous scoring. SANE models with $a_* = 0.94$, $R_{\rm low} = 1$, 10, and $R_{\rm high} = 10$ are ruled out by simultaneous scoring, but score fairly well in joint scoring. For the favored MAD models, when $R_{\rm low} = 1$, there are more images which simultaneously satisfy all constraints, but when $R_{\rm low} = 10$, the distributions generally stay closer to the observed data ranges and are thus favored by the joint scoring method. Due to differences between simultaneous and joint scoring results, we consider the inferred parameters of $R_{\rm low}$, $R_{\rm high}$, and a_* from passing models to be less robust than the overall trend that MAD models are favored in both scoring methods.

The simultaneous scoring method has the advantages of conceptual simplicity, and it accounts for correlations between the scoring metrics since it applies each constraint simultaneously. Simultaneous scoring is more strict and rules out more models than joint scoring, but it may be more limited by the finite number of images generated per model. The joint scoring procedure has the advantage of being more conservative in disfavoring models, but assumes the observational constraints are independent in calculating a joint likelihood. Instead, they are correlated (in particular $|m|_{\rm net}$, $\langle |m| \rangle$, and $|\beta_2|$ all show correlations).

The number of images in each model that passes each constraint individually (used in joint scoring) or simultaneously (used in simultaneous scoring) can be found in Appendix D of Event Horizon Telescope Collaboration et al. (2021b).

4.9.5 Combined EHTC V and current polarimetric constraints

EHTC V presented constraints on the GRMHD simulation models based on fits to the EHT total intensity data, model self-consistency (requiring a radiative efficiency less than that of

a thin accretion disk at the same black hole spin), and M87's measured jet power (requiring a simulation to produce a jet power consistent with a conservative lower limit of that from M87, $> 10^{42}\,\mathrm{erg\,s^{-1}}$). Those constraints ruled out MAD $a_* = -0.94$ models (from failing to satisfy the EHT image morphology), SANE models with $a_* = -0.5$, and all models with $a_* = 0$ (from failing to produce enough jet power). Here we retain only the jet power constraint, which is the most constraining and straightforward to apply to the expanded image library considered in this work.

Relativistic jets launched in GRMHD simulations (defined here as in EHTC V, with a cutoff of $\beta\gamma > 1$) are fully consistent with being produced via the Blandford–Znajek process (e.g., McKinney & Gammie, 2004; McKinney, 2006). As a result, $a_* = 0$ models have small or zero jet power, $P_{\rm jet}$, and are rejected by this constraint. These models can still produce significant total outflow powers ($P_{\rm out}$ in EHTC V) in a mildly relativistic jet or wind. Many other models with low values of $R_{\rm high}$ or moderate black hole spin are also ruled out by the jet power constraint. Combining the simultaneous scoring polarimetric constraints with the jet power constraint results in 15 remaining viable models: all MADs and all spanning the full range of non-zero a_* explored. This conclusion does not depend on the choice of the simultaneous or joint model scoring procedure.

4.10 Distributions of theory metrics for each model

Prograde, SANE models show rapidly decreasing $|m|_{\rm net}$ with increasing $R_{\rm high}$ and are significantly depolarized when $R_{\rm high} > 10$. This behavior was previously demonstrated by Mościbrodzka et al. (2017). The accretion flow electron temperature decreases with increasing $R_{\rm high}$, increasing the strength of Faraday rotation while also concentrating the emission at high latitudes behind the black hole (see also EHTC V). The emission is then depolarized when traveling through the Faraday-thick midplane plasma.

Retrograde SANE models, however, show nearly the opposite behavior, with depolariza-

tion maximized for $R_{\rm high} = 1$. At larger values of $R_{\rm high}$, linearly polarized emission appears on the near side of the midplane, producing coherent linear polarization structure that is not Faraday-depolarized.

MAD models at all spins show a mild degree of depolarization with increasing $R_{\rm high}$. The accretion flow electron temperature remains high even for large values of $R_{\rm high}$, since much of the plasma has $\beta \simeq 1$.

Similar qualitative behavior is seen in $\langle |m| \rangle$ and the amplitude of β_2. However, those quantities show less time variability (narrower distributions) than is seen in $|m|_{\rm net}$. As a result, observed ranges of those values are more constraining. In particular, the MAD models show consistent offsets where $\langle |m| \rangle$ and $|\beta_2|$ are lower for $R_{\rm low} = 10$ than $R_{\rm low} = 1$ models. Some spin dependence is also apparent, with high prograde spin usually corresponding to the highest degrees of ordered polarization.

When the β_2 amplitude is not strongly suppressed (e.g., by Faraday rotation), the β_2 phase distributions are related to intrinsic magnetic field structure (e.g., Figure 4.10 and Palumbo et al. 2020). Prograde spin, $R_{\rm high} = 1$ SANE models and retrograde spin, large $R_{\rm high}$ SANE models both show radial EVPA patterns, resulting in β_2 phase distributions near zero. MAD models show spin-dependent β_2 phase distributions for low values of $R_{\rm high}$, ranging from spiral patterns ($\arg \beta_2 \simeq -90\,{\rm deg}$) for retrograde spin to more radial patterns at high prograde spin. The patterns are relatively constant functions of $R_{\rm high}$ and $R_{\rm low}$, although with some shift of MAD prograde distributions to twistier EVPA patterns, particularly for $R_{\rm low} = 10$.

4.11 Discussion

The resolved EHT 2017 linear polarization map of M87 shows a predominantly azimuthal linear polarization (EVPA) pattern and relatively low fractional polarization of $\lesssim 20\%$ on $20\,\mu$as scales. We interpret the low fractional polarization as the result of Faraday rotation

internal to the emission region, which acts to rotate, scramble, and depolarize the resolved polarized emission. Adopting this constraint in a one zone model, we estimate typical values of particle density $n_{\rm e}$, magnetic field strength B, and electron temperature $T_{\rm e}$. In semi-analytic emission models with externally imposed, idealized magnetic field configurations, azimuthally dominated EVPA patterns are produced by poloidal (radial and/or vertical) magnetic field components. To fully capture the complicated combined effects from magnetic field structure, turbulence, relativity, and Faraday rotation on polarimetric images of M87, we turn to radiative transfer calculations from GRMHD simulations.

We compared a large image library of emission models from GRMHD simulations with metrics designed to capture these salient features of the data. The combined constraints of a predominantly azimuthal EVPA pattern and a low but non-zero fractional polarization are inconsistent with most SANE GRMHD models with weaker horizon scale magnetic fields. Some MAD models with relatively cold electrons, realized in our library by larger values of $R_{\rm high}$ and/or $R_{\rm low}$, remain consistent with the data. Here we discuss the implications of our results, and limitations in our set of theoretical models that may impact our interpretation.

4.11.1 Near horizon plasma and magnetic field properties in passing models

Both our one-zone and GRMHD models find similar plasma conditions in the 230 GHz emission region, driven by the requirements of weak 230 GHz absorption and strong 230 GHz Faraday rotation. In viable GRMHD models, we find average, intensity-weighted plasma properties in the emission region of $n_e \sim 10^{4-5}\,{\rm cm}^{-3}$, $B \simeq 7-30$ G, and $\theta_e \sim 8-60$. These are in good agreement with our one zone estimates (Section 4.7.1). We have also calculated the intensity-weighted values of the absorption and Faraday optical depth, τ_I and τ_{ρ_V}, over snapshots that simultaneously satisfy all our observational constraints. The median values are $\tau_I \simeq 0.1$ and $\tau_{\rho V} \simeq 50$. All of our viable images have $\tau_{\rho V} > 2\pi$, while 2 out of 73 have $\tau_I \gtrsim 1$, consistent with our assumptions in Section 4.7.1 that the plasma Faraday depth is

large while the Stokes $\mathcal{I}$ optical depth is small.

By quantitatively evaluating a large library of images based on GRMHD models (Section 4.9), we identify 25 out of 120 models that remain viable after applying constraints based only on EHT and ALMA-only polarimetric observations. Additionally applying a cut on jet power of $P_{\rm jet} > 10^{42}\,{\rm erg\,s^{-1}}$ (EHTC V) rules out the five viable SANE models and all $a_* = 0$ models. The precise number and identity of the viable models depends mildly on the chosen scoring procedure and on the Gaussian blurring kernel size applied to the EHT image reconstructions and library simulated images. The overall preference for MAD over SANE models is found from both the simultaneous and joint scoring procedures, as well as other variants. After applying the jet power constraint, no viable SANE models remain for any of the scoring methods we explored.

MAD models are associated with dynamically important magnetic fields. The significant poloidal components of those fields can produce a predominantly azimuthal polarization pattern (Figure 4.11), similar to those seen in idealized models with prescribed poloidal magnetic fields (Figure 4.10). Strong Faraday effects complicate a direct interpretation of the observed EHT polarization map in terms of those idealized models. Still, our more detailed comparison favoring MADs suggests the presence of dynamically important magnetic fields in the emission region on event-horizon scales.

In Figure 4.17 we present mass accretion rate and jet power distributions both for the viable models identified in EHTC V and when adopting the new constraints from polarimetry. Polarimetric constraints break degeneracies present in the single epoch total intensity data, allowing us to estimate a mass accretion rate onto the black hole of $\dot{M} \simeq (4-16)\times 10^{-4} M_\odot\,{\rm yr^{-1}}$. This corresponds to $\dot{m} = \dot{M}/\dot{M}_{\rm Edd} \simeq (3-12)\times 10^{-6}$, where $\dot{M}_{\rm Edd}$ is the Eddington accretion rate.[3] The measured radiative efficiency $\epsilon = L/\dot{M}c^2$ (where L is the bolometric luminosity) of the passing models is relatively high for a hot accretion

[3]The Eddington rate is defined as $\dot{M}_{\rm Edd} = L_{\rm Edd}/\epsilon_{\rm Edd}c^2$, where $L_{\rm Edd} = 4\pi GMm_pc/\sigma_T$ is the Eddington luminosity and we adopt an efficiency factor $\epsilon_{\rm Edd} = 0.1$. Note that this assumed efficiency factor $\epsilon_{\rm Edd}$ is distinct from the reported radiative efficiency $\epsilon = L/\dot{M}c^2$ measured from the simulations.

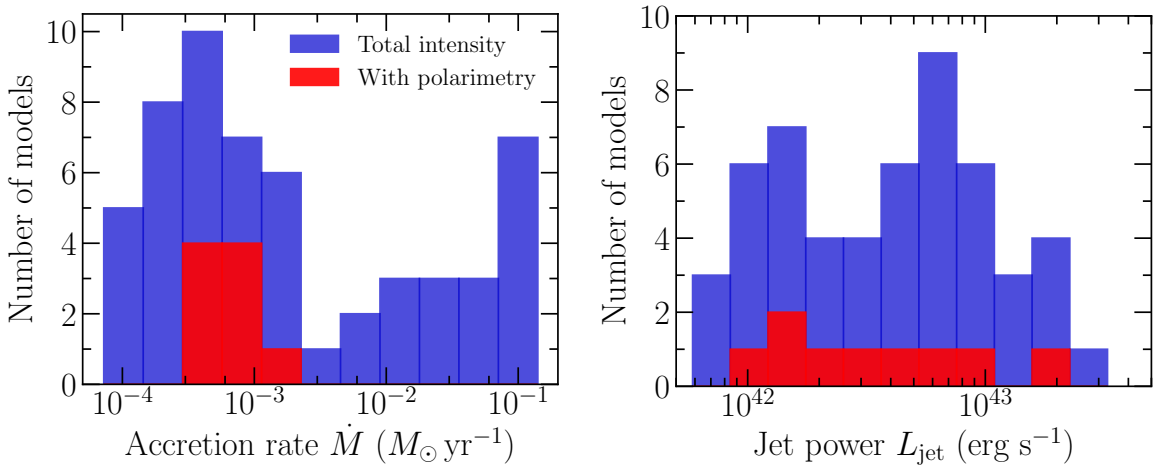


Figure 4.17: Average mass accretion rate (left) and jet power (right) for viable GRMHD models of M87 identified by selecting on total intensity data and jet power (blue), and when including polarimetric constraints from simultaneous scoring (red). We estimate a mass accretion rate of $\dot{M} \simeq (4-16)\times10^{-4} M_\odot\,\mathrm{yr}^{-1}$, resulting in a radiative efficiency $\epsilon \lesssim 1\%$ (cf. EHTC V). The jet powers produced by our models are $\sim 10^{42}-10^{43}$ erg s^{-1}, and the jet efficiencies are $\simeq 5-80\%$. Compared to EHTC V, the range of jet powers remains the same while the mass accretion rate is better constrained.

flow model: $\epsilon \lesssim 1\%$. These models have jet powers of $P_{\rm jet} \simeq 10^{42-43}\,\mathrm{erg\,s^{-1}}$.

The mass accretion rate found here is much lower than the Bondi rate calculated from *Chandra* observations (Di Matteo et al., 2003, see also Russell et al. 2015) and higher than that found from hybrid disk+jet models of the M87 SED (Prieto et al., 2016). Our inferred jet powers of $\lesssim 10^{43}\,\mathrm{erg\,s^{-1}}$ are at the lower end of the observed range. In particular, the jet power measured at the location of *HST-1* is $\sim 10^{43-44}\,\mathrm{erg\,s^{-1}}$ (Stawarz et al., 2006), and LOFAR observations suggest a jet power of $\sim 10^{44}$ erg s^{-1} was necessary within the last ∼million years to inflate the observed radio lobes on scales of ~ 80 kpc (de Gasperin et al., 2012).

Measurements of the accretion rate and the radiative efficiency can begin to constrain the microphysical plasma processes that heat electrons in M87, for example by inferring the fraction of the dissipated energy in the system that heats electrons, $\delta_{\rm e}$. In axisymmetric, self-similar, hot accretion flow models, a system with $\dot{M} \sim 10^{-5}\dot{M}_{\rm Edd}$ and a radiative efficiency $\epsilon \lesssim 1\%$ has a value of $\delta_{\rm e}$ in the range $0.1-0.5$ (see Figure 2 of Yuan & Narayan 2014). This

range is consistent with the range produced by simulations of turbulence and reconnection in the $\beta \sim 1$ regime (e.g., Rowan et al., 2017; Werner et al., 2018; Kawazura et al., 2019). Future studies using simulations with self-consistent electron heating and radiative cooling (Section 4.11.3) can better constrain δ_e and its dependence on local plasma parameters throughout the accretion flow and jet-launching region.

We have assumed that all effects responsible for the appearance of the EHT polarized image of M87 are captured within the relatively small GRMHD simulation spatial domain, $\lesssim 10^{2-3} r_g$. Goddi et al. (2021) developed a two-component model for the ALMA and image-integrated EHT data where each component is Faraday rotated by a different screen. The model demonstrates that the rotation measure of the compact component is unconstrained by the ALMA measurements alone, since the ALMA measurements are also sensitive to the Faraday rotation properties of the larger-scale component. In addition, the observed time variability in ALMA data (e.g., the RM sign change) can be explained by the observed EVPA variation of the compact core seen by the EHT. To produce the observed variability requires an RM of $\approx -6 \times 10^5 \,\mathrm{rad\,m^{-2}}$. The ALMA data do not constrain the location or nature of this Faraday screen, except that it must be relatively close to the compact core, $r \lesssim 10^5 \, r_g$.

For our favored plasma parameters for M87, we expect substantial Faraday rotation measure internal to the emission region itself, $\tau_{\rho_V} \gtrsim 2\pi$, consistent with that measured from viable GRMHD images. In a model of uniform, external Faraday rotation this Faraday depth at 230 GHz would correspond to an RM of $\lesssim 10^6 \,\mathrm{rad\,m^{-2}}$. Figure 4.18 (and in more detail Appendix B of Event Horizon Telescope Collaboration et al. (2021b)) shows that the apparent RMs measured from our GRMHD images span a wide range, often comparable to or larger than that implied from the Goddi et al. (2021) two-component model ($\lesssim 10^6 \,\mathrm{rad\,m^{-2}}$). For the low inclination angle of M87, the apparent RM measured from GRMHD images is not a good tracer of the mass accretion rate (Mościbrodzka et al., 2017) and originates close to the emission region, well within the simulation domain (Ricarte et al. 2020 and Appendix B of Event Horizon Telescope Collaboration et al. 2021b). The RM inferred from

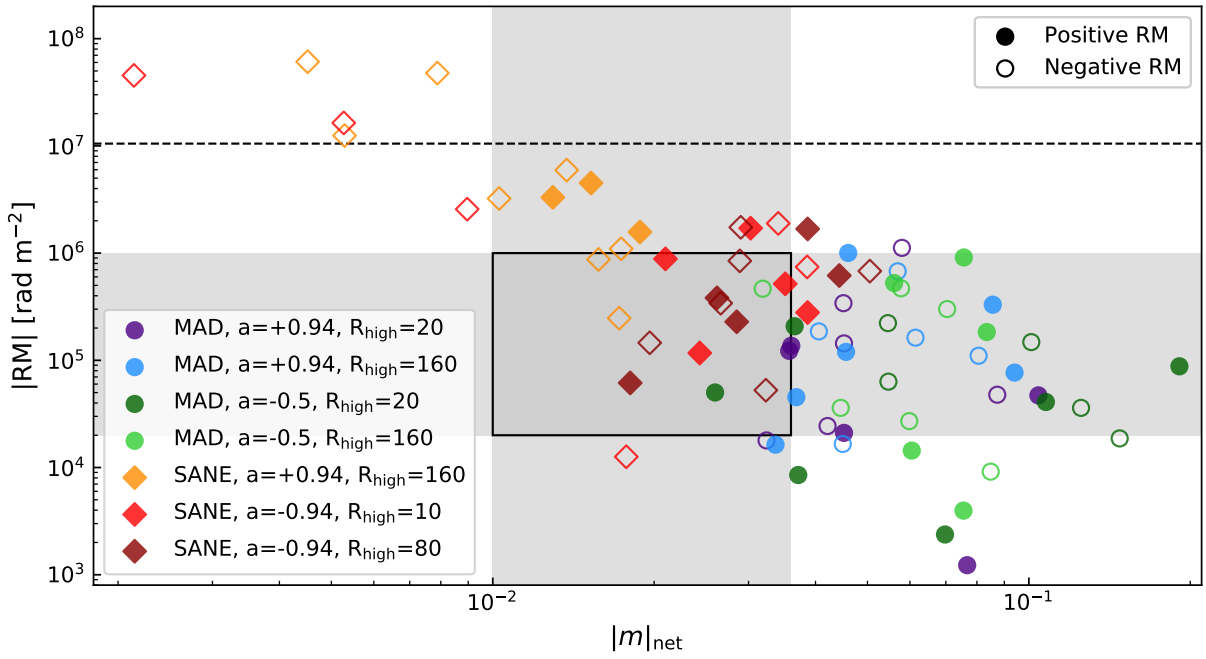


Figure 4.18: Absolute value of rotation measure (RM) versus net linear polarization $|m|_{\rm net}$ for a subset of our EHT GRMHD library models explored in more detail in Ricarte et al. (2020). Closed symbols represent positive RM while open symbols represent negative RM, revealing significant time variability across the 2500 GM/c^3 spanned by these snapshots. In grey, we plot our allowed region of $|m|_{\rm net}$ and bracket the range of core RM inferred from contemporaneous ALMA-only observations, $2-100\times10^4$ rad m^{-2} (Goddi et al., 2021). The dashed horizontal line demarcates the RM at which an EVPA rotation by π radians would have been observed between the 212 and 230 GHz frequency range used in the ALMA-only measurements, 1.05×10^7 rad m^{-2}. Despite large Faraday depths, a large fraction of these snapshots exhibit RMs consistent with simultaneous ALMA-only constraints. RM and $|m|_{\rm net}$ are anti-correlated, since larger Faraday depths lead to greater scrambling of the intrinsic polarization.

low-inclination GRMHD models of M87 can also vary rapidly and change signs (Ricarte et al., 2020), as seen in the ALMA-only data. As a result, the RM inferred from the two-component model in Goddi et al. (2021) is apparently consistent with the intrinsic properties of the GRMHD models studied here, without invoking an additional, external Faraday screen. At the same time, we cannot rule out that such an external screen could be present. Future EHT observations with wider frequency spacing can directly measure the resolved RM of the core and address this uncertainty.

Magnetic reconnection, MHD turbulence and collective plasma modes in collisionless hot accretion flows likely result in non-thermal particle acceleration. A provisional survey of the influence that a power law (i.e., non-thermal) electron distribution function has on the model scoring result can be found in Section 6.2 of Event Horizon Telescope Collaboration et al. (2021b). In brief, it is found that using a hybrid thermal-plus-powerlaw distribution function tends not to affect the structure of the EVPA map (β_2 amplitude and phase), but it may change the image-integrated and resolved linear polarization fractions. More realistic particle acceleration scenarios could be considered using *resistive* GRMHD simulations (e.g., Ripperda et al. 2019).

The above summary assumes that the mass accretion rate for the non-thermal models is $\sim$ the rate found in the thermal ones. If instead we compare the models at fixed flux density, the mass accretion rate of the hybrid model must be decreased. Generalizing the distribution function therefore introduces order unity uncertainties in the inferred mass accretion rate, radiative efficiency, and jet power. The changes in the polarimetric observables in a given snapshot are also larger at fixed flux density. For example, the full analysis, which appears in EHTC VIII, found that in the MAD $a_* = -0.5$ model, the $|m|_{\rm net}$ increases from 4.7% to 6% when adding non-thermal electrons.

4.11.2 Coherently polarized forward jet emission

As noted above, some SANE retrograde model images in the library show coherently polarized features even when the Faraday depth through the entire emission region is large. The observed polarized flux in those cases originates on the near side of the midplane and is not scrambled from Faraday rotation along the line of sight. A similar effect might be possible if non-thermal electrons could be accelerated efficiently in the low-density, strongly magnetized funnel region in front of the black hole.

It is beyond the scope of this paper to evaluate whether or how such a model might be realized physically, e.g., if some process could fill the funnel with high-energy electrons efficiently

enough to produce the observed 230 GHz luminosity from the funnel alone. EHTC VIII performed a cursory analysis of a single snapshot from a prograde $a_* = 0.94$ by assigning a non-thermal energy density $u_{\rm nth} = \alpha u_{\rm mag}$ wherever the magnetization $\sigma > 1$, where $\alpha = 0.02$ was the fraction of the magnetic energy density that is put into non-thermal particles (a pure power law distribution with $\gamma_{\rm min} = 100$ and $p = 3$).

In the purely thermal case, Faraday rotation depolarizes the emission at the EHT beam scale, producing low fractional polarization across the image that is inconsistent with EHT observations of M87. Adding power law electrons in the funnel produces coherent linearly polarized emission. If $u_{\rm nth} \propto$ the magnetic energy density, the power law emission is concentrated close to the black hole and lensed into a ring (Dexter et al., 2012). The weak forward jet component is strongly polarized but lies inside the observed ring, and it is thus potentially inconsistent with the EHT total intensity and polarimetric image. Although the augmented model was found to be consistent with constraints on the net and image-average linear polarization fraction, it is inconsistent with the observed β_2 phase of the M87 image.

In the example, the plasma was assumed to be composed of protons and electrons rather than e^+/e^- pairs. The latter are presumably more likely to form in the funnel (Mościbrodzka et al., 2011; Levinson & Cerutti, 2018; Chen et al., 2018; Crinquand et al., 2020; Wong et al., 2021b; Anantua et al., 2020a) and have different circular polarization properties. Future observations that constrain the resolved circular polarization structure might discriminate between pair and electron–ion plasmas in the emitting region. At longer wavelengths and larger scales, the limb-brightened jet structure of M87 (e.g., Walker et al., 2018) also suggests that the radiating electrons are not concentrated inside the funnel as modeled here.

4.11.3 Radiative models

Our GRMHD images use the parameterization of Mościbrodzka et al. (2016) to model the different electron and ion temperatures. The functional form of the temperature ratio, Equation 4.27, captures the general behavior seen in many simulations of electron heating in tur-

bulent or reconnecting collisionless plasmas; namely, the electron heating is more efficient (and thus the temperature ratio closer to unity) when $\beta < 1$ (e.g., Howes, 2010; Rowan et al., 2017). However, the actual distribution of T_e in a hot accretion flow reflects the balance of heating, cooling, and advection of hot electrons throughout the system. Even if the dissipation sets the initial electron temperature locally following Equation 4.27, the final temperature ratio can be different as electrons are advected and cool over time. Furthermore, the GRMHD simulations in the library considered here do not include radiative cooling. Our passing models for M87 favor a radiative efficiency of $\epsilon \sim 1\%$, however, and we may begin to worry if cooling is dynamically important in M87.

To assess these uncertainties, it will be useful to compare the results in this work with results from simulations performed with radiative GRMHD code. These codes typically use either the M1 closure method (e.g., Sądowski et al., 2013b; McKinney et al., 2014; Sądowski & Gaspari, 2017) or a Monte Carlo approach (e.g., Ryan et al., 2015; Ryan & Dolence, 2020) to track radiation and its interactions with the plasma near the black hole. In addition to the effects of cooling on the total gas, these codes can also evolve the electron temperature under the influence of cooling and different prescriptions for the uncertain electron heating physics (e.g., Ressler et al., 2015, 2017; Chael et al., 2018a; Ryan et al., 2018; Chael et al., 2019; Dexter et al., 2020).

We have checked the trends observed in the simulation library against the radiation GRMHD models of M87 presented in Ryan et al. (2018) (SANE) and Chael et al. (2019) (MAD). In both sets of simulations, the distribution of the $T_{\rm e}/T_{\rm i}$ on average qualitatively tracks the $R_{\rm high}$ prescription (Mościbrodzka et al., 2016), with values of effective $R_{\rm high} \sim 10 - 20$. We find similar effective $R_{\rm high}$ values even when the underlying electron heating prescriptions are quite different, as in the reconnection and turbulence-heated models used in Chael et al. (2019). In both cases, radiative cooling lowers the electron temperature in the funnel region so that the temperature ratio in the funnel $R_{\rm low} > 1$.

4.12 Predictions

We have identified a subset of a large parameter space of GRMHD models that is consistent with constraints derived from current EHT total intensity and polarimetric observations of M87. The models that pass our constraints on the polarimetric structure and jet power from M87 are all magnetically arrested (MAD) accretion flows. Here we make predictions for testing our interpretation with future observations.

4.12.1 Repeated observations

Repeated EHT observations of M87 at 230 GHz will continue to constrain the model parameter space. Figure 4.19 shows the time evolution of β_2 amplitude and phase for 200 snapshots of 3 viable library models: MAD $a_* = -0.5$, $R_{\rm low} = 10$, $R_{\rm high} = 20$; MAD $a_* = +0.5$, $R_{\rm low} = 10$, $R_{\rm high} = 80$; and MAD $a_* = +0.94$, $R_{\rm low} = 10$, $R_{\rm high} = 80$. The observer inclination in these models was 17 deg or 163 deg for the retrograde and prograde models, respectively.

Both quantities show variations on timescales from days to months. The phase and amplitude of β_2 should change over the course of a week of observations. In EHTC VII, we observe changes in the the β_2 amplitude and phase over the week of observations in 2017, and use the results from two epochs to define our acceptable parameter ranges. Figure 4.19 suggests that occasionally the observed changes in β_2 on $\sim$week timescales can be much more dramatic than we observe in 2017, with variations in β_2 phase of 90 deg on short timescales for some models.

The scatter in both quantities on longer $\gtrsim$ month timescales is much larger than the uncertainty range derived from the EHT2017 measurements. If our passing GRMHD models accurately describe the 230 GHz emitting region in M87, future EHT observations should detect variability in the polarization structure. According to current models, the time-averaged β_2 amplitude and $\langle |m| \rangle$ should remain similar to the current values for prograde

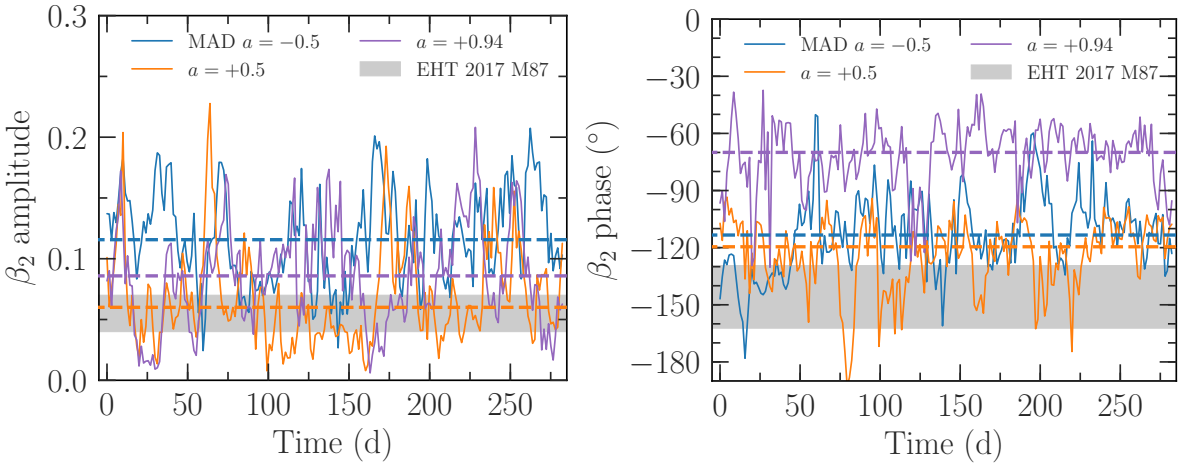


Figure 4.19: Amplitude (left) and phase (right) of β_2 as a function of time for three viable GRMHD library models identified here (points, all with $R_{\rm low} = 10$) compared to ranges measured from EHT 2017 M87 data (gray shaded region). The dashed lines show the median values for each model. The retrograde spin model predicts higher β_2 amplitude in future observations. In the high prograde spin model, the median β_2 phase is closer to zero than the observed range in 2017. Changes in both quantities occur on timescales of weeks to months, and should be apparent in future EHT data sets.

spin models and tend toward larger values for retrograde spin models. For high prograde spin (or many SANE models), the β_2 phase should on average be closer to zero than we observe in 2017.

4.12.2 Future observations at 260 and 345 GHz

In selecting models, we have focused on metrics corresponding to salient features of the data. We have not attempted to compare models in detail to specific features of the reconstructed polarimetric images, most notably the apparently depolarized bright patch in the eastern part of the image. We do note that such depolarized features occur in many of our library images, particularly in MAD models with $R_{\rm low} = 10$. If the eastern patch in the 2017 image is depolarized due to Faraday rotation, it may be possible to tell with future higher frequency observations. In addition to internal Faraday rotation, the sense of the EVPA pattern may also be subject to a net, coherent rotation due to external Faraday rotation. At higher frequency, Faraday rotation is suppressed and EHT observations will see the intrinsic

magnetic field pattern more clearly.

The full analysis in EHTC VIII presents a preliminary study of the library image polarization metrics at higher frequencies and found that $|m|_{\rm net}$ and $\langle|m|\rangle$ are expected to increase with frequency while the net circular polarization $|v|_{\rm net}$ remains small and nearly constant. In the preliminary test models, the EVPA pattern did not change significantly with frequency, suggesting that the net EVPA pattern is due to magnetic field structure rather than coherent Faraday rotation. This hypothesis can be tested with future multi-frequency observations.

4.13 Conclusions

The EHT has produced resolved polarized intensity maps in the near-horizon region around the supermassive black hole in M87. Taken together with image-integrated data from simultaneous observations with ALMA, these images constrain the space of accretion flow and jet models used to interpret the EHT total intensity image with broad implications for jet launching near a black hole event horizon. Here we summarize the main results of the analysis.

- We interpret the depolarization seen in EHT images as the result of beam depolarization due to Faraday rotation internal to the emission region ($\tau_{\rho V} \sim 50$). In the context of one-zone models and combined with the size and brightness temperature of the total intensity image, we estimate an average emission region plasma density of $n_e \sim 10^{4-7}\,{\rm cm}^{-3}$, magnetic field strength of $B \sim (1-30)$ G, and $T_e = (1-12)\times 10^{10}$ K.

- The net EVPA pattern of the M87 polarization maps is predominantly azimuthal. In the context of semi-analytic models with imposed, idealized magnetic field geometry, such a pattern can be reproduced using a significant component of poloidal (radial and/or vertical) magnetic field. The presence of such magnetic fields in a rotating fluid would imply that the magnetic fields are dynamically important. However, significant

Faraday rotation may be present, and it is not clear whether the observed EVPA pattern can be interpreted in terms of magnetic field structure alone.

- To capture the effects of realistic magnetic field structure, plasma conditions, and Faraday rotation and conversion, we have compared salient observables to a large library of simulated polarimetric black hole images. Our comparison metrics include net circular polarization, net and image-averaged linear polarization, and the β_2 coefficient of Palumbo et al. (2020). We find that β_2 is the most constraining metric.

- The model scoring procedures disfavor most models from the GRMHD image library from polarimetric observations alone. Many weakly magnetized (SANE) models are too depolarized or show an EVPA pattern that is too radial. Many strongly magnetized (MAD) models are too coherently polarized. The polarization fraction is generally set by the Faraday rotation depth close to the emission region. MAD models more frequently produce azimuthal EVPA patterns, as expected for magnetic field structures that include a significant poloidal field component. Combined with a conservative lower limit on the jet power of M87, only strongly magnetized (MAD) models remain viable. We use those remaining models to estimate the mass accretion rate onto the central supermassive black hole to be $\dot{M} = (4 - 16) \times 10^{-4}\, M_\odot\, \mathrm{yr}^{-1}$. The average plasma parameters found from GRMHD images are in good agreement with those inferred from one zone models.

- The favored models show time variability in the polarization metrics we have considered. The median values found at several epochs should be sufficiently well measured to distinguish between the current retrograde and prograde spin models. At higher frequencies of 260 and 345 GHz, weaker Faraday effects should result in an increased degree of polarization. Continued imaging with the EHT and advances in radiative and non-thermal theoretical models will further constrain the electron distribution and magnetic field structure in the jet-launching region.

Chapter 5

Mass entrainment at the jet–disk boundary

Magnetic fields lines are trapped in black hole event horizons by accreting plasma. If the trapped field lines are lightly loaded with plasma, then their motion is controlled by their footpoints on the horizon and thus by the spin of the black hole. In this chapter, I investigate the boundary layer between the lightly loaded field lines near the pole and a dense, equatorial accretion flow. I present an analytic model for aligned prograde and retrograde accretion systems and argue that there is significant shear across this "jet–disk boundary" at most radii for all black hole spins. The analytic model predicts the strongest shear in the retrograde aligned accretion scenario, so I use numerical simulations of retrograde accretion to show instability of the jet–disk boundary in that case. I find that the mixing layer episodically loads plasma onto trapped field lines where it is heated, forced to rotate with the hole, and permitted to escape outward into the jet. By using Lagrangian tracer particles to follow the flow of matter through a particular simulation, I find that the time-averaged mass-loading rate is $\sim 1\%$ the accretion rate through the event horizon.

5.1 Introduction

According to Alfvén's theorem, magnetic fields lines are frozen into highly conducting plasmas and are advected with the plasmas as they move under the influence of external forces. This freeze-in effect operates near black holes when the accreting plasma falls onto the hole, and thus it is natural for a black hole to have field lines that thread its event horizon. If

This chapter is lightly adapted from G. N. Wong, *et al.*, ApJ, in press.

the horizon-threading field lines are open and lightly loaded with plasma so that the local magnetization[1] is much larger than unity

$$\sigma \equiv \frac{B^2}{\rho c^2} \gg 1 \tag{5.1}$$

in the region close to the horizon, then their motion is controlled by gravity, and they are forced to rotate if the black hole has nonzero spin.

Forced rotation of field lines was first studied by (Blandford & Znajek, 1977, hereafter BZ) who solved a force-free magnetosphere model in the limit that the black hole dimensionless spin $a_* \equiv Jc/(GM^2) \ll 1$ (here $J \equiv$ spin angular momentum and $M \equiv$ mass). BZ found that the field behaves as if it were anchored in a star rotating with frequency

$$\Omega_F \equiv \frac{1}{2}\Omega_H = \frac{a_*}{8}\frac{c^3}{GM} + O(a_*^3). \tag{5.2}$$

Here, Ω_H the rotation frequency of the event horizon. Field line rotation produces an outward-directed energy current at the horizon. In the force-free limit this is known as the BZ effect, whereas if the field lines are more heavily loaded it is also sometimes called the magnetohydrodynamic (MHD) Penrose process (Takahashi et al., 1990). The BZ effect is a favored mechanism for powering extragalactic radio jets.

In recent decades, numerical general relativistic magnetohydrodynamics (GRMHD) simulations have been used to study black hole accretion and the BZ mechanism (see Davis & Tchekhovskoy (2020) and Komissarov & Porth (2021) for reviews). In GRMHD models with a trapped magnetic flux Φ, a low-density region forms around an axis parallel to the accretion flow angular momentum vector as plasma falls down the field lines into the hole or is expelled to larger radius. This low-density region, with $\sigma \gg 1$, contains horizon-threading field lines moving with rotation frequency Ω_F and an associated, outward-directed energy

[1]Here, B is the strength of the magnetic field, ρ is the rest-mass density of the plasma, and c is the speed of light. In this paper, we use Lorentz–Heaviside units for electromagnetic quantities.

current (Poynting flux; McKinney & Gammie, 2004). In what follows we will refer to this region as the *jet*. It is difficult for numerical codes to robustly evolve parts of the simulation domain with low density and high σ, like in the jet, so semi-analytic magnetosphere models are often invoked to study these regions (see, e.g., Ogihara et al., 2021).

The jet is bounded by an accretion flow that pins magnetic flux in the hole. We will refer to the accretion flow as a *disk*, although it may have sub-Keplerian rotation. At the boundary layer between the jet and the disk, the density contrast is large. The plasma velocity can also change dramatically, with maximal shear occurring when the black hole and disk rotate in opposite directions (a *retrograde* disk).

The jet–disk boundary layer has large shear and strong currents. It can suffer instabilities that lead to mass loading onto the jet's open field lines. It may also be an important particle acceleration site (see the reviews of Ostrowski, 1999; Rieger, 2019, for particle acceleration in relativistic shear layers). This paper considers the jet–disk boundary layer in the relativistic regime, within $\sim 20\ GM/c^2$ of the event horizon.

In Section 5.2 we provide simple estimates for shear at the jet–disk boundary layer. In Section 5.3 we describe the GRMHD simulations we use to study the jet–disk boundary layer, and in Section 5.4, we explore the dynamics of the boundary layer by using tracer particles to both analyze the flow of matter through state space and investigate mass loading into the jet. Along the way we discuss the disk structure for retrograde accretion. In Section 5.5 we consider model limitations, convergence, and possible extensions. Section 5.6 provides a summary and a guide to the main results.

5.2 Scaling and estimates

We now define the physical parameters that describe accretion systems, identify their ranges for the systems we consider, and provide an analytic estimate for flow dynamics at the jet–disk boundary layer.

5.2.1 Parameters

We consider radiatively inefficient accretion flows (RIAFs; Reynolds et al. 1996) where radiative cooling is negligible, motivated by Event Horizon Telescope (EHT) observations of M87* and Sgr A*, which have accretion rates $\dot{m} \equiv \dot{M}/\dot{M}_{\rm Edd} \ll 1$ ($\dot{M}_{\rm Edd}$ is the Eddington accretion rate) and are therefore near or in this regime. RIAFs are geometrically thick disks, with ratio of scale height H to local radius R of order 1.

In general, the angular momentum of accreting matter far from the horizon may be tilted with respect to the black hole's spin angular momentum. Although there are plausible scenarios that produce zero tilt, there is at present no way of rejecting models with strong or even maximal (180°) tilt. In this paper we restrict attention to systems where the orbital angular momentum of the accreting plasma is parallel or anti-parallel to the black hole spin vector (prograde or zero tilt and retrograde or maximal tilt, respectively). Disks with intermediate tilt are a subject of ongoing study (Fragile et al., 2007; McKinney et al., 2013; Morales Teixeira et al., 2014; Liska et al., 2018; White et al., 2019).

In addition to a_*, $\dot{m}$, and tilt, black hole accretion flows are characterized by Φ, the trapped magnetic flux measured through the contour formed by the black hole's equator. Accretion of flux with a consistent sign eventually increases $|\Phi|$ until the accumulated magnetic flux is large enough that magnetic pressure $B^2 \sim (\Phi/(GM/c^2)^2)^2$ balances accretion ram pressure ρc^2. Since $\dot{M} \sim \rho c (GM/c^2)^2$, when the dimensionless flux $\phi \equiv \Phi/\sqrt{G^2M^2\dot{M}/c^3}$ approaches a critical value $\phi_c \sim 15$ (Tchekhovskoy et al. 2011, but we use the normalization of Porth et al. 2019), the field can push aside infalling plasma and escape.

The unstable equilibrium with $\phi \sim \phi_c$ is known as a magnetically arrested disk (MAD; see Bisnovatyi-Kogan & Ruzmaikin, 1974; Igumenshchev et al., 2003; Narayan et al., 2003), in contrast to accretion flows with $\phi \ll \phi_c$, which are said to follow standard and normal evolution (SANE; see Narayan et al., 2012; Sądowski et al., 2013a). Notice that ϕ is determined by the nonlinear evolution of the flow and is not trivially related to the initial conditions, although initial conditions have been identified that lead to SANE or MAD outcomes over

finite integration times. We will consider both SANE and MAD accretion flows.

5.2.2 Shear at the jet–disk boundary

Changes in velocity across the jet–disk boundary may drive Kelvin-Helmholtz instability. What is the expected velocity difference? The jet and disk are unsteady and strongly nonaxisymmetric in the numerical GRMHD models that motivate this calculation. In the interest of producing a model that can be studied analytically, we nevertheless treat the system as axisymmetric and steady, and because this is already a drastic approximation, we use a nonrelativistic fluid model for simplicity.

The jet can be idealized as a steady flow anchored in an object rotating with angular velocity Ω_F. For a steady, axisymmetric, nonrelativistic MHD wind with plasma angular velocity Ω and generalized specific angular momentum L, angular velocity changes with cylindrical radius R like

$$\Omega = \Omega_F \frac{1}{1+M_{\rm A}^2} + \frac{L}{R^2}\frac{M_{\rm A}^2}{1+M_{\rm A}^2} \tag{5.3}$$

(e.g., Ogilvie, 2016), where $M_{\rm A}^2 \equiv v_p^2/v_{\rm A}^2$ is the Alfvén Mach number, defined as the ratio of the poloidal plasma velocity to the Alfvén velocity $\boldsymbol{v}_{\rm A} = \mathbf{B}/\sqrt{\rho}$. Since $\sigma \gg 1$, evidently $v_{\rm A} \simeq c$.

Particles flow inward at the horizon and outward at large radius, and therefore a steady state can be achieved only if plasma is loaded onto field lines at intermediate radius. We assume this occurs, perhaps through turbulent diffusion or through pair production (in numerical GRMHD models plasma is added via numerical floors; see Wong et al. 2021b for a study of drizzle pair production in this region), and that there is a stagnation point at $r \sim$ few $\times\, GM/c^2$ between an inner, inflow Alfvén point ($M_{\rm A}^2 = 1$) and an outer, outflow Alfvén point. The outer Alfvén point is close to the light cylinder $r_l \sin\theta = c/\Omega_F$.

Equation 5.3 implies that for $M_{\rm A}^2 \ll 1$, $\Omega \sim \Omega_F$, and for $M_{\rm A}^2 \gg 1$ the specific angular momentum of the wind is conserved. Inside of the light cylinder, in the limit that $a_* \ll 1$,

rotation is controlled by the rotation frequency of the hole Ω_H, like $\Omega_F \approx \Omega_H/2 \approx a_*/8$, so

$$\Omega \approx \begin{cases} \dfrac{\Omega_H}{2} & r < r_l \\ \dfrac{2c^2}{\Omega_H}\dfrac{1}{(r\sin\theta)^2} & r > r_l \end{cases} \tag{5.4}$$

The jet–disk boundary is at $\theta_{\rm JD}$, so the outer light cylinder radius is $r_l = (8/a_*)(GM/c^2)/(\sin\theta_{\rm JD}) + \mathcal{O}(a_*)$. Taking $\sin\theta_{\rm JD} \simeq 1/\sqrt{2}$, then $r_l \simeq (11/a)(GM/c^2)$.

The disk rotates with approximately constant angular velocity $\Omega = s\Omega_K$ on spherical surfaces; here, $\Omega_K = (GM)^{1/2}r^{-3/2}$ is the Keplerian angular velocity and $0 < s < 1$ measures how sub-Keplerian the accretion flow is. Numerical simulations suggest $s \lesssim 1/2$ for MADs and ~ 1 for SANEs.

The toroidal component of the velocity difference across the jet–disk boundary is thus

$$\Delta v_\phi \simeq r\sin\theta_{\rm JD}(\Omega_F - s\Omega_K). \tag{5.5}$$

Without a model for flow along the field lines it is not possible to constrain the other components of the velocity difference. For retrograde accretion with $a_* < 0$, the two angular frequencies in Equation 5.5 have the same sign and the magnitude of the velocity jump is at least of order the orbital speed. The velocity difference is approximately c at $r = r_l$. For prograde accretion with $a_* > 0$, the shear vanishes at $r = 4(s/a_*)^{2/3}(GM/c^2)$, and as in the retrograde case, the velocity difference is $\sim c$ at $r = r_l$.

5.2.3 Stability of the jet–disk boundary

The jet–disk boundary is associated with sharp changes in density and magnetic field. The jet contains a laminar $\sigma > 1$ plasma, analogous to a pulsar wind, that rotates with the black hole. The disk contains a turbulent $P_{\rm gas}/B^2 \sim 1$ plasma whose angular momentum need not be related to the spin of the central hole. The relative orientation of the shear, jet magnetic field, and disk magnetic field may vary as turbulence in the disk produces varying conditions

at the boundary.

Is the jet–disk boundary linearly stable? If we model the boundary layer as an infinitely thin current-vortex sheet, then we expect to capture the main features of the linear theory; finite thickness H tends to suppress instability for modes with wavelengths smaller than or of order H, and the fastest growth is at wavelengths $\sim H$. The current-vortex sheet can be subject to Kelvin–Helmholtz instability (KHI) as well as the plasmoid instability (Loureiro et al. 2007). High-resolution axisymmetric models of black hole accretion flows (Ripperda et al. 2020; Nathanail et al. 2020) see evidence for plasmoid instability at the jet–disk boundary, but we do not, perhaps due to inadequate resolution. We therefore focus on KHI. It is well known that magnetic fields weaken the KHI because they resist corrugation of the vortex sheet. Do magnetic fields stabilize the jet–disk boundary?

A general linear theory of the plane-parallel, *relativistic*, ideal current-vortex sheet does not exist. Osmanov et al. (2008) consider the special case where magnetic field is oriented parallel to the velocity shear and the density, pressure, and field strength are continuous across the sheet. They do not consider the large density contrast that is an important feature of the jet–disk boundary problem.

The linear theory of the plane-parallel, compressible, *nonrelativistic*, ideal current-vortex sheet is better understood. The general (arbitrary field orientation on either side of the sheet) incompressible case was considered by AXFORD (1960); Shivamoggi (1981) considers aligned and transverse fields; Sen (1964) and Fejer (1964) consider a general, arbitrarily oriented field on either side of the sheet. The stability of a finite-width layer has been considered in a well-known analysis by Miura & Pritchett (1982), but an analytic dispersion relation is not available. Since the general, nonrelativistic problem is relatively tractable we provide a brief discussion and use it to obtain a qualitative understanding of stability of the jet–disk boundary.

Consider a plane-parallel, nonrelativistic, current-vortex sheet. The flow velocity and magnetic field are constant away from the sheet, which we position at $z = 0$. Let $i = J$

denote the low-density (jet) side and $i = D$ the high-density (disk) side. In equilibrium, $v_{\mathrm{A}z}$ vanishes and total pressure is continuous across $z = 0$.

Now consider a perturbation of the form $f(z)\exp(ik_x x + ik_y y + i\omega t)$ with $f(z) = \exp(\kappa z)$, where κ is in general complex. The general dispersion relation is

$$\lambda_J m_D + \lambda_D m_J = 0 \tag{5.6}$$

$$\lambda_i = \rho_i \left[(\omega - \boldsymbol{k}\cdot\boldsymbol{v}_i)^2 - (\boldsymbol{v}_{\mathrm{A}i}\cdot\boldsymbol{k})^2\right] \tag{5.7}$$

$$m_i = \sqrt{k^2 + \frac{(\omega - \boldsymbol{k}\cdot\boldsymbol{v}_i)^4}{{c_{si}}^2(\boldsymbol{v}_{\mathrm{A}i}\cdot\boldsymbol{k})^2 - c_{si}^2(\omega - \boldsymbol{k}\cdot\boldsymbol{v}_i)^2}} \tag{5.8}$$

(Sen, 1964; Fejer, 1964). Here, $c_s \equiv$ sound speed, $c_m^2 \equiv v_{\mathrm{A}}^2 + c_s^2$ is the magnetosonic speed, and $\boldsymbol{v}$ is the plasma velocity. The exponential factor κ can be m_i or $-m_i$ (see Equation 5.8) depending on the boundary condition and whether $z > 0$ or $z < 0$.

The general dispersion relation cannot be solved analytically. In the case of interest to us, however, $\rho_J \ll \rho_D$, $c_{sD} \sim v_{\mathrm{A}D}$, and $c_{sJ} \sim c_{sD}$. Furthermore, physics provides a hint to the mathematical solution: the field in the jet is stiff (the Alfvén speed is large due to the low density), motivating us to look for instability in modes with $\boldsymbol{k}\cdot\boldsymbol{v}_{\mathrm{A}J} = 0$. This is enough to make analytic progress. Taking $\rho_J/\rho_D \sim \epsilon^2 \ll 1$ and assuming that $\boldsymbol{k}\cdot\boldsymbol{v}_{\mathrm{A}D} \sim \epsilon$, we can solve the dispersion relation to lowest order in ϵ. The relevant mode has

$$\omega^2 = (\boldsymbol{k}\cdot\boldsymbol{v}_{\mathrm{A}D})^2 - \frac{\rho_J}{\rho_D}\left[\boldsymbol{k}\cdot(\boldsymbol{v}_J - \boldsymbol{v}_D)\right]^2, \tag{5.9}$$

which suggests that the current-vortex sheet is unstable when $\boldsymbol{k}\cdot\boldsymbol{v}_{\mathrm{A}D}$ is sufficiently small, which we have confirmed by numerically solving the full dispersion relation.

In Equation (5.9) the nonrelativistic current-vortex sheet is unstable for small ρ_J. This is precisely the limit where one might worry about relativistic corrections: if $B_J^2/\rho_J > 1$, then the inertia of the jet is dominated by the magnetic field. In a fully relativistic analysis (Y. Du et al., in preparation), the current-vortex sheet has a near-identical dispersion relation in the

limit $\rho_J \to 0$, except that ρ_J/ρ_D in the above dispersion relation is replaced by B_J^2/ρ_D.

Evidently the current-vortex sheet is not generically unstable at large density contrast: a particular configuration of magnetic fields is needed for instability. The disk contains a turbulent magnetic field that is constantly changing strength and orientation, while the jet has a steadier field. This suggests a picture in which turbulent mixing driven by the KHI is episodic and occurs when jet and disk magnetic fields are aligned or anti-aligned. Mixing as a result of nonlinear development of the KHI will then only occur when there exist modes with growth times that are small compared to the correlation time of the turbulent eddies.

5.2.4 Dissipation at the jet–disk boundary

The jet–disk boundary would appear to be a fertile setting for particle acceleration: particles that cross the boundary from the disk plasma frame to the jet plasma frame gain energy in a process akin to Fermi acceleration. This has been investigated by, e.g., Berezhko & Krymskii (1981); Jokipii & Morfill (1990); Ostrowski (1990) (see Rieger 2019 for a review), usually in the context of extragalactic radio jets that are kiloparsecs from the central source. Sironi et al. (2021) performed 2D particle-in-cell simulations of the shear layer between a relativistic, magnetically dominated electron–positron jet and a weakly magnetized ion–electron plasma and showed that the nonlinear evolution of Kelvin–Helmholtz instabilities leads to magnetic reconnection, which can in turn drive particle acceleration. The formation of magnetic islands at the jet–disk boundary (see, e.g., Nathanail et al., 2020; Ripperda et al., 2020) can also lead to particle acceleration; this process has been extensively investigated in kinetic simulations of current sheets.

To schematically address this question, we adopt a turbulent resistivity model for dissipation in the jet–disk boundary with magnetic diffusivity $\eta \simeq \alpha\, W\, \Delta v$, where α is the inverse of the magnetic Reynolds number, the width of the boundary layer is $W \sim fR$ ($f < 1$; here, $R \equiv$ cylindrical radius) and $\Delta v \sim c$, so that $\eta \simeq \alpha f c R$. Next, we assume that the boundary is steady, axisymmetric, and follows $R = R_0(z/z_0)^\beta$, with the jet intersecting the

horizon at (R_0, z_0). We assume that the magnetic flux in the jet $\Phi \simeq \pi B R^2$ is approximately independent of R and thus take $B \simeq \Phi (z/z_0)^{-2\beta}/(\pi {R_0}^2)$.

If the magnetic field in the disk is similar in magnitude to that in the jet but randomly oriented, the dissipation rate per unit volume in the boundary layer is $\Lambda \sim \alpha B^2 (c/(fR))$, and the total dissipated power per unit height z is independent of f:

$$\frac{dP}{dz} = \frac{1}{\pi} \alpha c \frac{\Phi^2}{R_0^3} \left(\frac{z}{z_0} \right)^{-3\beta} \left(1 + \beta^2 \frac{R_0^2}{z_0^2} \left(\frac{z}{z_0} \right)^{-2+2\beta} \right)^{1/2}. \tag{5.10}$$

Notice that this scales asymptotically as $z^{-1-2\beta}$ for $\beta \geq 1$, so nearly all dissipation occurs close to the black hole. Integrating over z, the dissipated power is

$$P = \frac{\alpha c^5 \Phi^2}{\pi (GM)^2} F(\beta, z_0/R_0), \tag{5.11}$$

where F is a dimensionless function of order unity. The power differs only by a factor of a_*^2/α from the Blandford–Znajek power (e.g., Tchekhovskoy et al. 2011). To sum up: a fraction $\sim \alpha/a_*^2$ of the jet power can be dissipated in the jet–disk boundary close to the black hole; this provides additional motivation for a numerical study.

5.3 Simulating black hole accretion

We now study the jet–disk boundary layer using GRMHD simulations.

5.3.1 Numerical setup

We integrate the equations of GRMHD using the `iharm` code, a descendant of the second order conservative shock capturing scheme HARM (Gammie et al., 2003). Written in a

coordinate basis, the governing equations of GRMHD are

$$\partial_t \left(\sqrt{-g}\rho_0 u^t\right) = -\partial_i \left(\sqrt{-g}\rho_0 u^i\right), \tag{5.12}$$

$$\partial_t \left(\sqrt{-g}T^t{}_\nu\right) = -\partial_i \left(\sqrt{-g}T^i{}_\nu\right) + \sqrt{-g}T^\kappa{}_\lambda \Gamma^\lambda{}_{\nu\kappa}, \tag{5.13}$$

$$\partial_t \left(\sqrt{-g}B^i\right) = -\partial_j \left[\sqrt{-g}\left(b^j u^i - b^i u^j\right)\right], \tag{5.14}$$

$$\partial_i \left(\sqrt{-g}B^i\right) = 0, \tag{5.15}$$

where the plasma is defined by its rest-mass density ρ_0, its four velocity u^μ, and b^μ is the magnetic field four-vector following McKinney & Gammie (2004). Here, $g \equiv \det(g_{\mu\nu})$ is the determinant of the covariant metric, Γ is a Christoffel symbol, and i and j denote spatial coordinates. In Equations 5.14 and 5.15, we express components of the electromagnetic field tensor $F^{\mu\nu}$ as $B^i \equiv {}^\star F^{it}$ for notational simplicity. The stress–energy tensor $T^\mu{}_\nu$ contains contributions from both the fluid and the electromagnetic field:

$$\begin{aligned} T^\mu_\nu = &\left(\rho_0 + u + P + b^\lambda b_\lambda\right) u^\mu u_\nu \\ &+ \left(P + \frac{b^\lambda b_\lambda}{2}\right) g^\mu_\nu - b^\mu b_\nu, \end{aligned} \tag{5.16}$$

where u is the internal energy of the fluid and the fluid pressure P is related to its internal energy through an adiabatic index $\hat{\gamma}$ with $P \equiv (\hat{\gamma} - 1)\, u$. The `iharm` code has been extensively tested and converges at second order on smooth flows (Gammie et al., 2003). A comparison of contemporary GRMHD codes can be found in Porth et al. (2019).

Our model has several limitations. First, we treat the accreting plasma as a nonradiating ideal fluid of protons and electrons. We do not consider effects due to anisotropy and conduction (Sharma et al. (2006); Johnson & Quataert (2007), but see Foucart et al. (2017) for an evaluation of the limits of this approximation). We also neglect radiation. This approximation may be inappropriate in systems with high mass accretion rates, like M87 (Dibi et al., 2012; Ryan et al., 2017), but it is sensible in systems with low $\dot{m}$ like Sgr A*

(but see Yoon et al. 2020, who show a different result under the assumption that the ions and electrons are perfectly coupled). The equations of nonradiative GRMHD are invariant under rescalings of both length and density, and so our numerical results can be scaled to the desired M and $\dot{M}$.

The `iharm` code evolves plasma on a logically Cartesian grid. For these simulations, we use FMKS coordinates, which are a modified version of the conventional horizon-penetrating Kerr–Schild (KS) coordinates. We use outflow boundary conditions for the radial direction, and we use a reflecting boundary condition at poles that mirrors the elevation components of the magnetic field and fluid velocity across the one-dimensional border.

We have added a passive tracer particle capability to `iharm` to track mass loading into the jet. Each tracer particle is introduced with probability proportional to the coordinate particle density $\sqrt{-g}\rho u^t$, where ρ is the rest-mass density, g is the determinant of the covariant metric, and u^t is the time component of the four velocity. Initial positions are uniformly distributed in the coordinate basis in each zone. Particles are advected with the fluid according to

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{u^i}{u^t} \tag{5.17}$$

where x^i are the spatial components of the tracer particle's position and u^μ is the fluid four velocity.

The computational cost of evolving the tracer particles alongside the fluid scales linearly with the number of particles; we use $\approx 2^{25}$ particles, and this noticeably increases simulation cost. We therefore use completed GRMHD simulations to identify an epoch of interest, restart the fluid simulation at the beginning of the epoch, initialize the particles, and re-evolve the fluid to the end of the epoch.

The `iharm` code has several limitations. It is not robust when $\sigma \gg 1$ (e.g. in the strong cylindrical explosion test in Komissarov, 1999) or when the ratio of the gas pressure to the magnetic pressure $\beta \equiv 2P_{\rm gas}/B^2 \ll 1$. Numerical stability is ensured by imposing artificial

ceilings on σ and $1/\beta$ in each zone at each time step, which are enforced by resetting the density or internal energy density to a floor value that depends on position but not on time. This has a minimal effect on the flow (as can be checked by varying the ceilings), but it does inject matter into the nearly evacuated funnel region, where σ is large and β is small.

The fluid sector is initialized with a perturbed Fishbone–Moncrief torus solution (Fishbone & Moncrief, 1976), which is parameterized by the inner disk edge radius $r_{\rm in}$ and pressure maximum radius $r_{\rm max}$. The thermal energy is perturbed to seed the instabilities that jump start accretion (including the magnetorotational instability). The SANE models have $r_{\rm in} = 6\ M$ and $r_{\rm max} = 12\ M$ in a domain that extends from within the horizon to $r_{\rm out} = 50\ M$. The MAD models have $r_{\rm in} = 20\ M$ and $r_{\rm max} = 41M$ in a domain that extends to $r_{\rm out} = 1000M$. Our MAD disks are larger than our SANE disks. Figure 5.1 shows the initial conditions for plasma and magnetic field in representative SANE and MAD simulations.

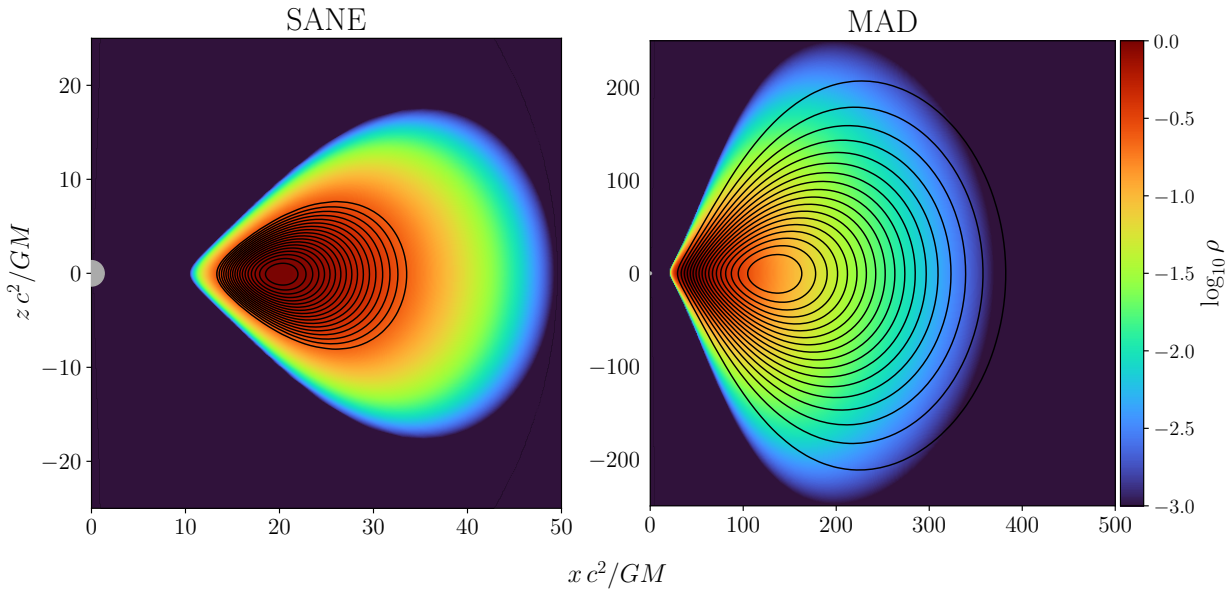


Figure 5.1: Initial distribution of plasma and magnetic field for representative retrograde SANE (left) and MAD (right) simulations. Both black holes have $a_* = -0.94$. The initial plasma density and magnetic field are axisymmetric. The central black hole is plotted at the center left of each panel. Color encodes $\log_{10}$ of plasma density. Magnetic field lines, which are purely poloidal, are overplotted in black. Notice that the domain of the MAD is 10x larger than the SANE simulation domain.

The initial magnetic field is described by the toroidal component of the vector four-potential $A_\phi(r,\theta)$. For SANE disks

$$A_\phi = \max\left[\frac{\rho}{\rho_{\max}} - 0.2, 0\right], \tag{5.18}$$

where $\rho_{\max}$ is the maximum initial plasma density. For MAD disks the initial field is concentrated toward the inner edge of the disk and forced to taper at large r according to

$$A_\phi = \max\left[\frac{\rho}{\rho_{\max}}\left(\frac{r}{r_0}\sin\theta\right)^3 e^{-r/400} - 0.2, 0\right], \tag{5.19}$$

where r_0 is chosen to be the inner boundary of the simulation domain (B. R. Ryan, priv. communication).

5.3.2 Simulations

Table 5.1 provides a summary of the models we consider. Our simulations are similar to the retrograde ones generated for the EHT simulation library in EHTC V, except that: our simulations are evolved twice as long to mitigate natural stochasticity in matter entrainment; and a subset of our simulations are rerun at multiple resolutions.

We focus on six retrograde simulations with $a_* = -0.5$ or -0.94. By convention, negative spins means that the black hole spin is anti-parallel to the angular momentum of the accretion flow (i.e. tilt is 180°). For each spin, we consider MAD and SANE models. We set the magnetic flux (and thus MAD or SANE state) by varying the field structure in the initial conditions.

Each simulation was run for at least 20,000 GM/c^3 and has an initial transient phase during which the initial torus relaxes, and magnetic winding and a combination of the Rayleigh–Taylor and Kelvin–Helmholtz instabilities operates. The transient phase is followed at each radius by a turbulent quasi-equilibrium, with equilibrium radius, defined as the largest radius where $d\dot{M}/dr \simeq 0$, increasing as $r_{\rm eq} \sim t^{2/3}$ (see, e.g., Penna et al. 2010; Dexter

ID	Flux	a_*	$r_{\rm in}$	$r_{\rm max}$	$r_{\rm out}$	Resolution	Notes
Sa-0.5	SANE	-0.5	6	12	50	288x128x128	medium disk
Sa-0.94	SANE	-0.94	6	12	50	288x128x128	medium disk
Ma-0.5	MAD	-0.5	20	41	1000	384x192x192	large disk
Ma-0.94_192	MAD	-0.94	20	41	1000	192x96x96	large disk
Ma-0.94_288	MAD	-0.94	20	41	1000	288x128x128	large disk
Ma-0.94	MAD	-0.94	20	41	1000	384x192x192	see note†
Ma-0.94_448	MAD	-0.94	20	41	1000	448x224x224	large disk

Table 5.1: GRMHD Simulation Parameters. Retrograde GRMHD fluid simulations parameters. Flux labels the relative strength of the magnetic flux at the horizon, a_* describes the spin of the black hole, $r_{\rm in}$ and $r_{\rm max}$ are parameters for the initial Fishbone–Moncrief torus, $r_{\rm out}$ is the outer edge of the simulation domain, resolution gives the $N_r \times N_\theta \times N_\phi$ number of grid zones in the simulation. † Multiple realizations of the 384x192x192 MAD $a_* = -0.94$ large disk simulations were run. Part of one realization included tracer particles.

et al. 2020 for a discussion). Beyond $r_{\rm eq}$, the flow is strongly dependent on initial conditions, so we consider information only from $r < r_{\rm eq}$. GRMHD models may be in equilibrium at large radii near the poles if there are strong outflows and the outflow structure is independent of the structure of the surrounding unequilibrated disk.

Our MAD simulations are run with bulk fluid adiabatic index $\Gamma = 13/9$, and our SANE simulations are run with $\Gamma = 4/3$ to be in agreement with EHTC V and Porth et al. (2019).

5.4 Results

We begin by discussing characteristic differences between MAD and SANE accretion flows before considering each of our simulations in detail. We explore the properties of fluid flow at small radii and within the jet, and then we relate outbursts in the MAD flows to magnetic flux ejection events. We explore qualitative features of the jet–disk boundary layer, including the development of Kelvin–Helmholtz instability. Finally, we use tracer particles to study mass entrainment across the jet–disk boundary layer.

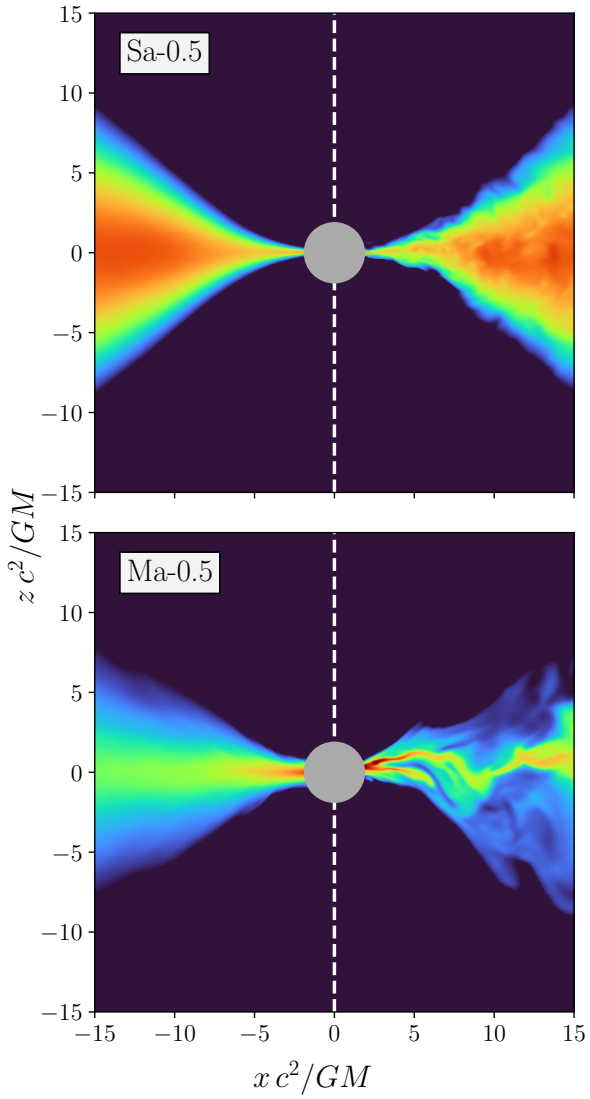


Figure 5.2: Logarithmic plots over three decades of density in the poloidal plane for $a_* = -0.5$ MAD and SANE models. Each image shows time- and azimuth-averaged density (left panels) and time slices at azimuth $\phi = 0$ (right panels). The density is particularly variable in the MAD models, where the time slice is not well approximated by the average state. The density is less variable in the SANE models, where the time slice and average state are comparatively similar.

5.4.1 Overview

It is convenient to divide low-luminosity black hole accretion flows into three regions: (1) the matter-dominated disk of plasma near the midplane, which on average flows inward, (2) the magnetically dominated, polar Poynting jet, and (3) the virial temperature intermediate

region, which contains the jet–disk boundary layer and the corona (here defined as the region with $\beta \sim 1$). In a region extending from the event horizon out to somewhat beyond the innermost stable circular orbit (ISCO), the inflow plunges supersonically onto the hole and fluctuates strongly. Notice that the jet we consider here (at horizon scales) is dynamically distinct from the jet at large radius.

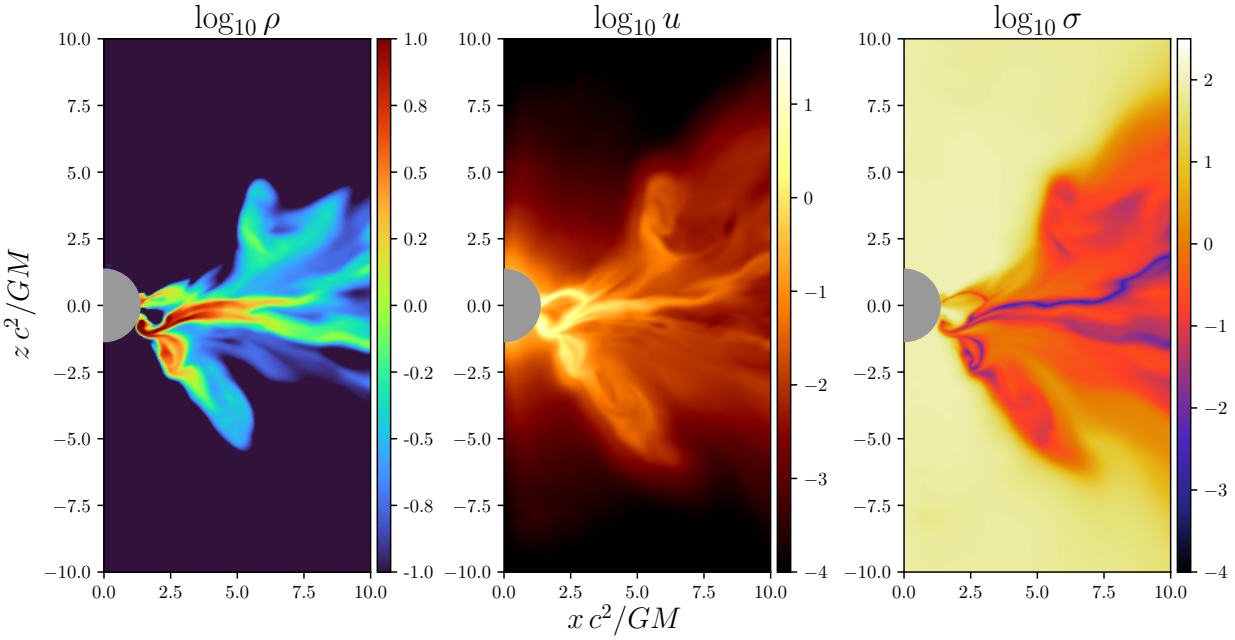


Figure 5.3: Azimuthal slice from an individual time slice of the $a_* = 0.94$ retrograde MAD simulation. Left panel: log density of plasma near the black hole. Center panel: log internal energy of the plasma $u = \rho T$. Right panel: plasma magnetization $\sigma = b^2/\rho$. The high σ, low-density conical regions around the poles are the jet funnel. The disk is the low σ, high-density region near the midplane. The intermediate region between the funnel and the disk with $\sigma \approx 1$ is the corona. The disordered accretion near the horizon is accentuated by streams of infalling plasma that are characteristic of MAD flows.

SANE and MAD accretion flows exhibit qualitatively different behavior. SANE models are relatively tame: plasma falls uniformly from the ISCO to the event horizon, the boundary of the accretion disk remains well defined, and the time-averaged accretion state is a fair approximation of an individual time slice. In contrast, MAD accretion is choppy and tends to proceed in isolated, thin plasma streams that begin far from the hole and plunge onto it. MAD accretion is punctuated by violent eruptions that release excess trapped magnetic

flux. Although the flux ejection events are not understood in detail, their structure suggests a Rayleigh-Taylor interaction between the disk and hole (see, e.g., Marshall et al., 2018). For MAD flows, the time average is often not a good approximation to a single time slice. These differences are particularly apparent in Figure 5.2, which shows log density for sample SANE and MAD models and compares the time-averaged solution (left) to representative time slices (right). In SANE models it is easy to separate the high-density disk from the low-density jet region. In contrast, in MAD models, identifying the location of the jet–disk boundary is a challenge.

In Figure 5.3, we show a typical time slice on a poloidal slice of an $a_* = -0.94$ MAD model, where the strength of the magnetic flux near the horizon prevents steady disk accretion. Here, accretion occurs when plasma streams break from the bulk disk at large radius and plunge onto the hole. These streams are not confined to the midplane as they fall. Figure 5.4 shows the projected locations of tracer particles in the same MAD $a_* = -0.94$ flow of Figure 5.3 but viewed from above. The color of each particle corresponds to the linear density of particles in a three-dimensional voxel of space centered at the particle and is used to visualize the complicated vertical structure of the flow. The figure shows one accretion stream connecting the disk and the hole in the bottom right, and the launch of two new streams in the upper right.

5.4.2 Counterrotation and the disk

As the black hole rotates, trapped magnetic field lines wind around the polar axis and produce a Poynting jet via the BZ mechanism. In the jet–disk boundary layer, however, the jet field lines (that rotate with the hole) are mixed with disk field lines (that rotate against the hole in retrograde models). This interaction leads to an exchange of angular momentum via magnetic and fluid stresses. Some of the infalling plasma then acquires negative u_ϕ, i.e., its specific angular momentum aligns with the black hole spin.

Exchange of angular momentum in the jet–disk boundary layer is more noticeable in

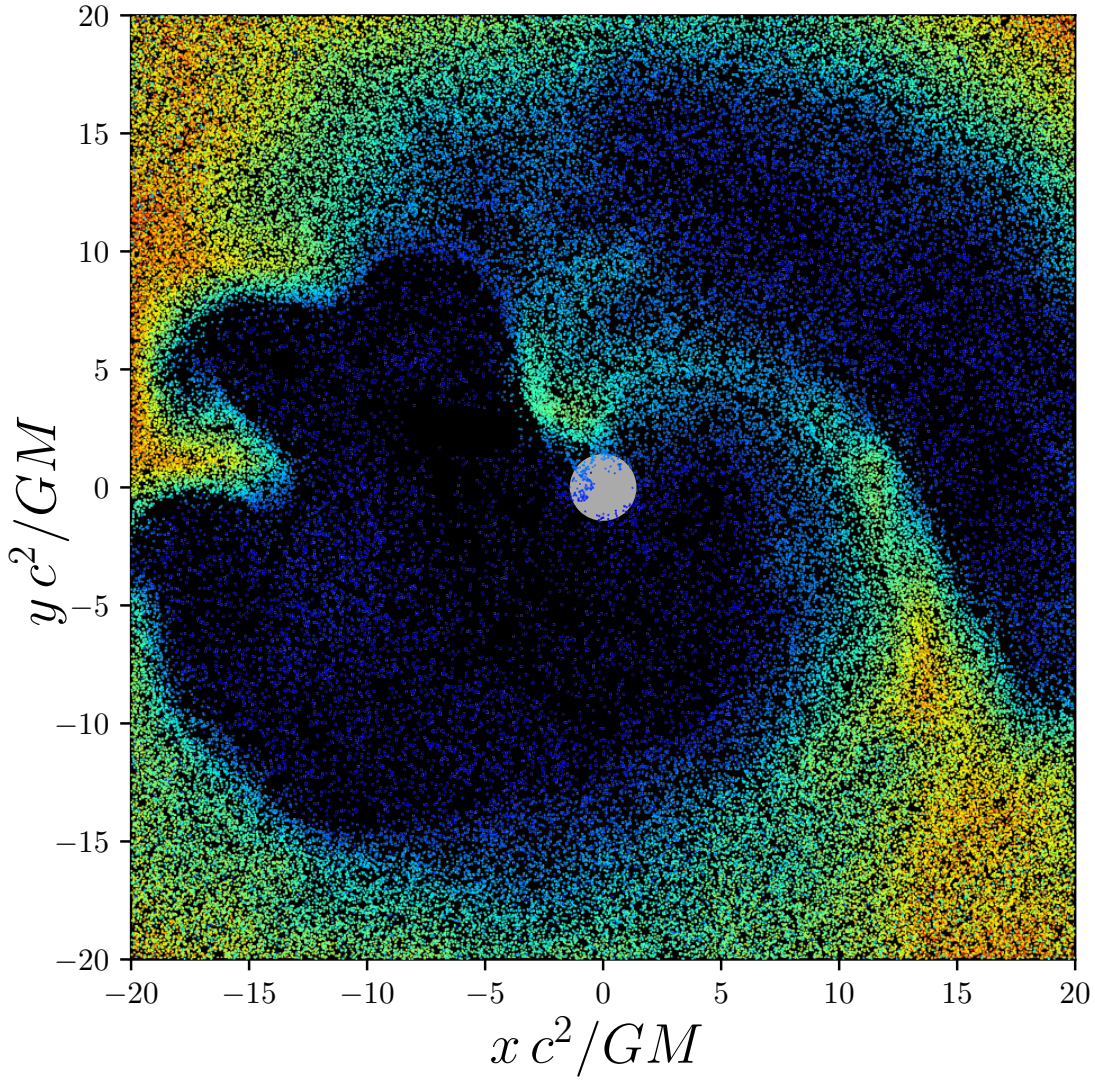


Figure 5.4: Tracer particle position for the MAD $a_* = -0.94$ model, projected onto the equatorial plane. Particle color varies linearly with local rest-mass density. The event horizon is a gray sphere. The inner region of the accretion flow is chaotic and characterized by plasma streams that break off the main disk at large radius. Plasma streams experience large magnetic torques (u_ϕ may change sign) as they plunge toward the horizon.

MAD models, where accretion occurs in streams and where the magnetic field tends to be stronger. In MAD models, the inhomogeneous flow magnifies the effects of magnetic torques, since some equator-crossing field lines are lightly loaded (in contrast to SANE models, in which the equator-crossing field lines pass through a dense disk). Moreover, the more concentrated magnetic flux tubes in the MAD models can result in stronger torques

(see Porth et al., 2021): when matter in the accretion stream with $u_\phi > 0$ interacts with a flux tube with $u_\phi < 0$, the plasma is rapidly braked and its angular momentum is reversed. Figure 5.5 shows an example of this interaction as counterrotating field lines collide with the corotating field lines near the horizon. During these events, the front edge of an accretion stream commonly erodes and accelerates radially outwards.

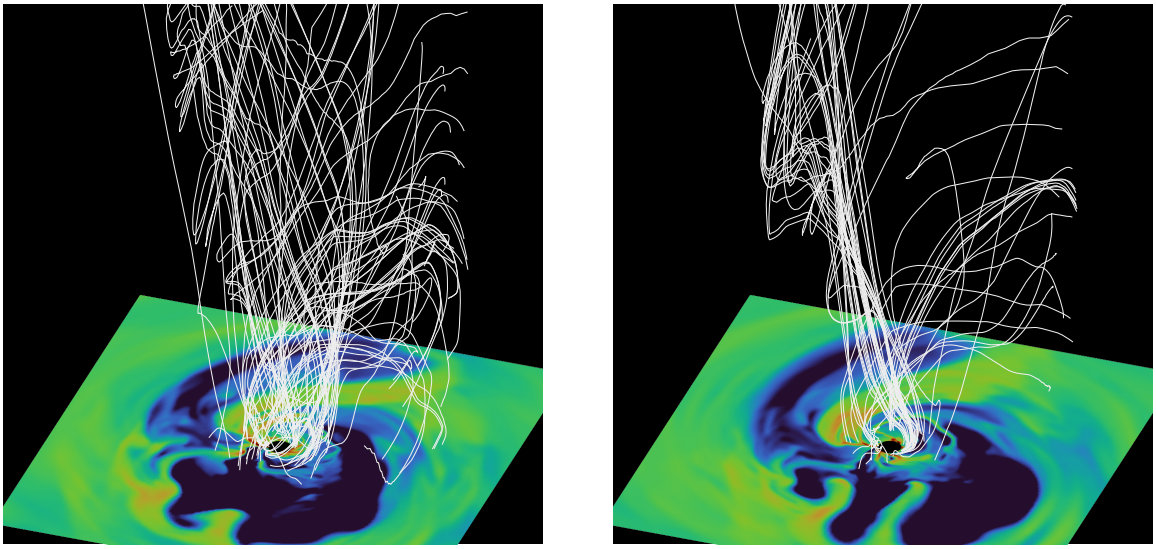


Figure 5.5: Interaction between disk and jet magnetic field lines. Magnetic field lines that intersect the disk at small radii are shown for two sequential time slices of the plasma evolution. Field lines are sampled according to magnetization in the midplane. The colored surface shows the logarithm over two decades of density in the midplane of the simulation, and the event horizon is plotted as a black circle in the center of the plane. Left panel: the same time slice as shown in Figure 5.4, rotated 45° counterclockwise. Magnetic field lines emanating from the high-density region toward the left of the figure trace an accretion stream and are disk dominated. Magnetic field lines that wind the opposite direction make up a flux tube and are being pulled clockwise with the hole as it spins. The two sets of field lines are about to collide. Right panel: same simulation approximately 50 GM/c^3 later. Disk-threading and funnel-threading magnetic field lines have interacted, and a much stronger flux tube passes through the midplane in the low-density region to the right of the hole.

The stronger angular momentum transfer in MAD flows produces more disorder in the inner region of the accretion flows. This difference between MAD and SANE models can be seen in Figure 5.6, which plots the time-integrated distributions of rest mass over u_ϕ, r and v^r, r. The infalling matter accelerates within the plunging region (close to the ISCO) in both

MAD and SANE flows, but the widths of the distribution of u_ϕ and v_r at a given radius differ sharply: the MAD models have larger width because they experience larger fluctuations.

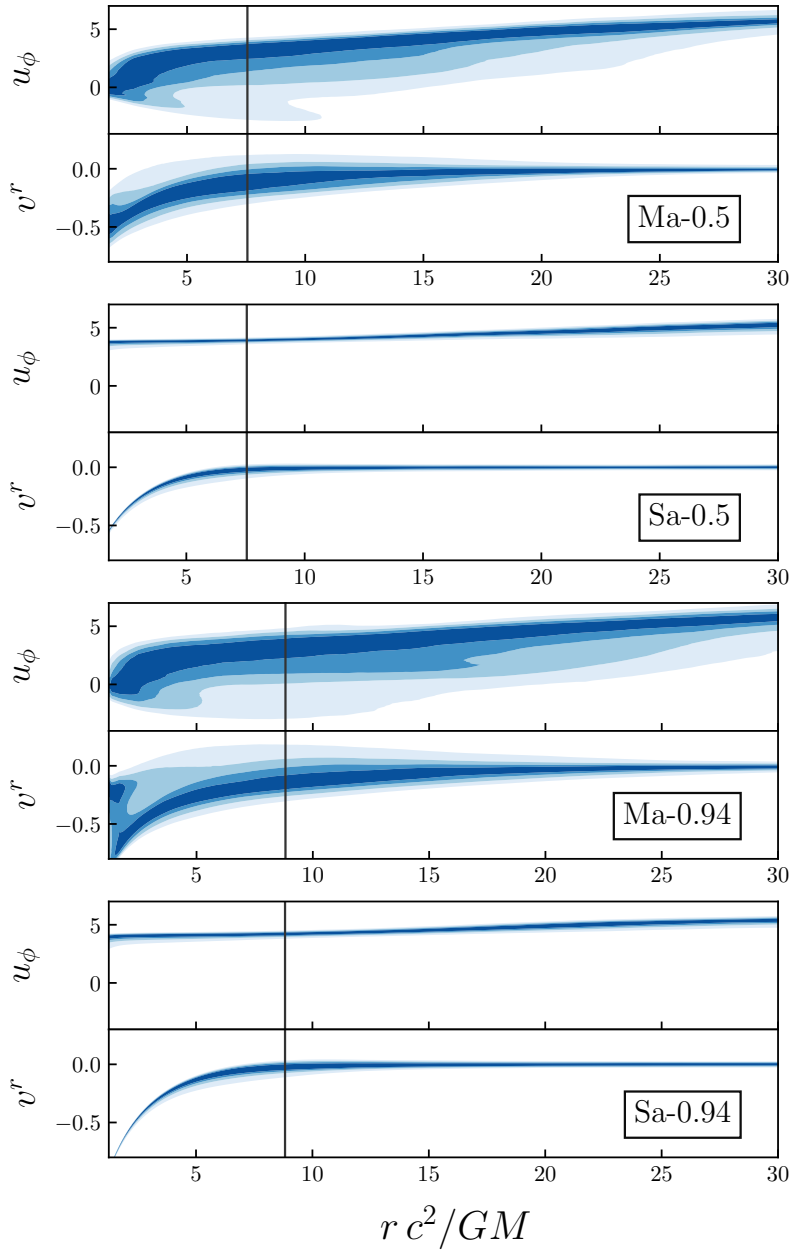


Figure 5.6: Distribution of matter in the angular momentum and radial velocity vs. radius ($u_\phi - r$ and $v^r - r$) planes for the four fiducial simulations. The vertical gray line marks the ISCO. The color scale is linear and shows the distribution of matter at each radius. In the SANE models the plasma lies on a well defined curve associated with Keplerian rotation as it accretes. In the MAD models plasma is perturbed away from the disk even before it enters the plunging region.

ID	a	b
Sa-0.5	0.22	1.8
Sa-0.94	0.18	1.8
Ma-0.5	0.07	2
Ma-0.94	0.1	2

Table 5.2: Funnel wall ($u_\phi = 0$ surface) fit parameters. Best fit parameters of the $z = Ax^b$ model for the location of the zero angular momentum surface in the GRMHD models.

5.4.3 Jet wall shape

In general, it is challenging to identify the jet–disk boundary since there is no clear criterion that distinguishes matter in the jet from matter in the disk (although proxy surfaces derived from magnetization or the Bernoulli parameter have been used in the past). Nevertheless, it is straightforward to find the surface where $u_\phi = 0$. Since u_ϕ has a definite sign in the jet, this surface may be a reasonable tracer of the boundary.

Figure 5.7 shows an azimuthal time slice of plasma density and angular momentum in the MAD $a_* = -0.94$ simulation and overplots the flow of the plasma. The lines change color at the $u_\phi = 0$ surface, which broadly separates outgoing matter from infalling matter. The extended jet–disk boundary is turbulent and mixes mass, angular momentum, and energy between the two regions. Figure 5.8 plots time- and azimuth-averaged u_ϕ for each of six models. We fit the $u_\phi = 0$ surface (within $r < 30\,GM/c^2$) to $z = ax^b$ and plot it as a dashed line. Recall that the boundary produced from the (ϕ, t)-averaged data may not be a good approximation to the boundary at fixed ϕ, t, especially for MAD models. The parameters for the fit are reported in Table 5.2.

In Figure 5.9 we plot $\langle u_\phi \rangle$, where the brackets indicate an average over time and azimuth versus elevation at four radii in each of the simulations. In MAD flows, we see that the average u_ϕ of matter in the midplane at $\theta = \pi/2$ decreases with radius; this makes sense since horizon-scale accretion flow is much choppier in MADs. The average u_ϕ of the plasma tends to increase with radius in both the disk and in the funnel. The point where u_ϕ changes sign corresponds to the location of the jet–disk boundary layer and roughly tracks the shape

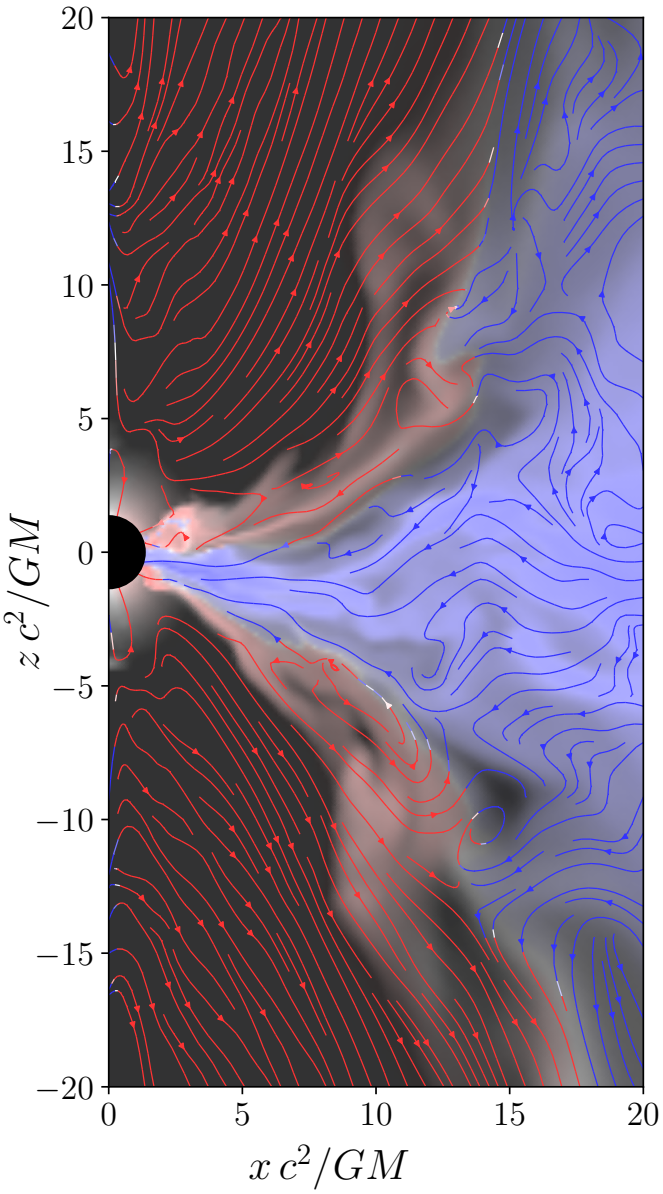


Figure 5.7: Time slice of a MAD, $a_* = -0.94$ model. Brightness shows plasma density, color saturation encodes value of u_ϕ, and flow lines describe the poloidal motion of the plasma. The jet–disk boundary is visible as the surface where u_ϕ changes sign. Eddies tend to form at the jet–disk boundary as infalling, positive u_ϕ matter interacts with outflowing, negative u_ϕ matter. The sign of u_ϕ in the funnel is set by the sign of black hole spin.

of the jet. In our SANE simulations, the boundary layer is resolved by $\gtrsim 16$ zones at all radii, and the jet spans approximately 10 zones at $r = 20\,GM/c^2$ and approximately 40 zones at $r = 2\,GM/c^2$. The boundary layer in our MAD simulations spans approximately $\gtrsim 30$ zones at all radii, and the jet is resolved by between 20 and 60 zones at $r = 20\,GM/c^2$

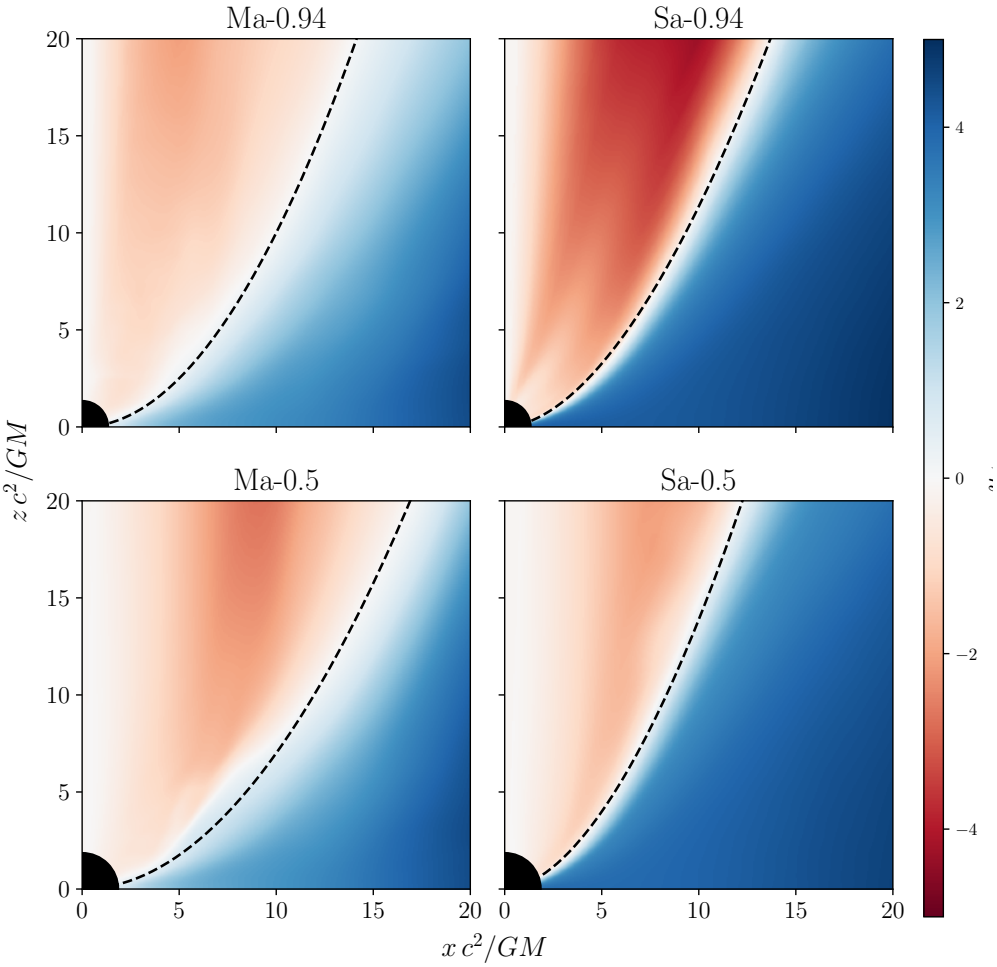


Figure 5.8: Density-weighted poloidal profile of u_ϕ for each of the four fiducial models after time and azimuthal averaging. The black circle at the origin marks the extent of the event horizon. All simulations have a similar structure: a parabolic jet (boundary defined by $u_\phi = 0$) and a peak in u_ϕ away from the pole.

and $r = 2\,GM/c^2$ respectively.

5.4.4 Mass entrainment

The shear layer at the jet–disk boundary is episodically unstable in our models. As instabilities develop, plasma from the disk is transported across the boundary, reverses direction, and is entrained into the jet. We use tracer particles to study mass entrainment and track matter that passes through the mixing region. The computational cost of tracking tracer

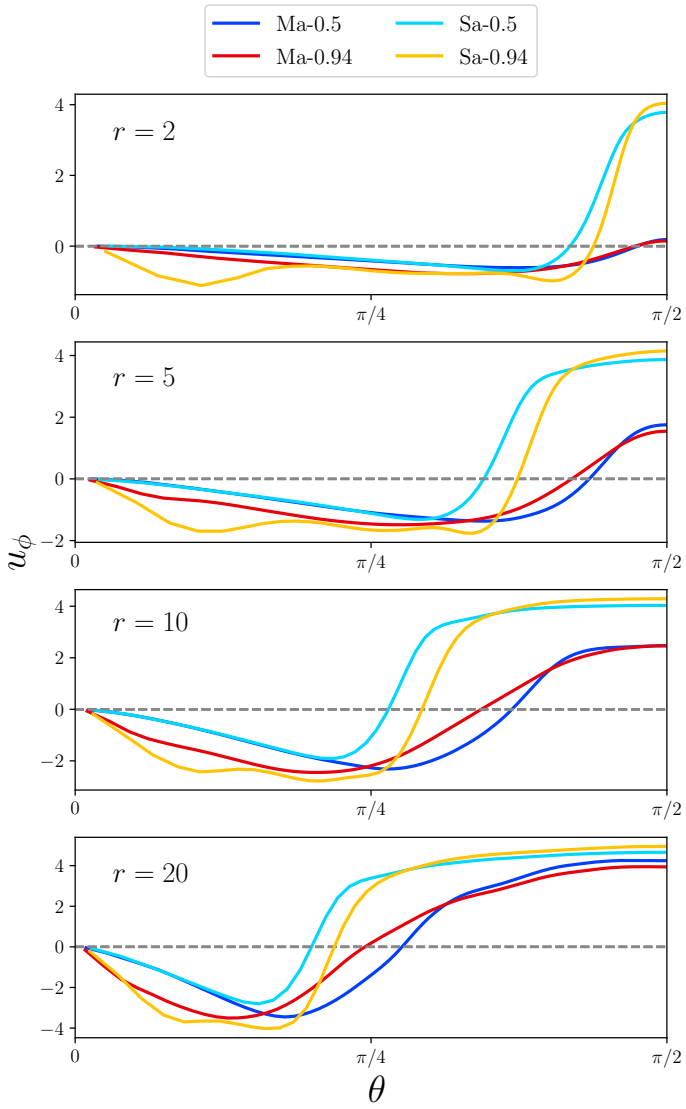


Figure 5.9: Profile of u_ϕ vs. elevation at $r = 2, 5, 10$, and 20 GM/c^2 for each of the models in Figure 5.8. Notice that $u_\phi < 0$ implies angular momentum aligned with the black hole. The average u_ϕ of plasma at small radii is smaller in MAD models than SANE models. The latitude of the shear layer within which u_ϕ changes sign increases with radius, corresponding to a narrowing jet. The (average) shear layer is wider for MAD models because their jet–disk boundaries fluctuate over a wider range in latitude. As matter flows out in the jet, magnetic torques increase u_ϕ.

particles in the global flow over the course of the entire simulation makes a full study prohibitively expensive. We instead perform a single high-resolution, high-cadence study that focuses on the evolution of approximately 3.2×10^6 particles within the inner region of the accretion flow over a $500\,GM/c^3$ interval. We chose to consider a range of time in the MAD

$a_* = -0.94$ model because it corresponded to an active period when multiple KHI knots are easily identifiable.

Entrained particles satisfy two criteria: they begin with $v^r < 0$ and $u_\phi > 0$, and they leave the simulation at the outer boundary with $u_\phi < 0$. In the mixing layer tracer particles may repeatedly transition between the disk and jet; we define entrainment to have happened for a tracer particle when its u_ϕ and v^r change sign for the last time. Because this definition of entrainment depends on the worldline of a fluid parcel, it is not immediately analogous to any quantity that can be directly computed from the raw fluid data.

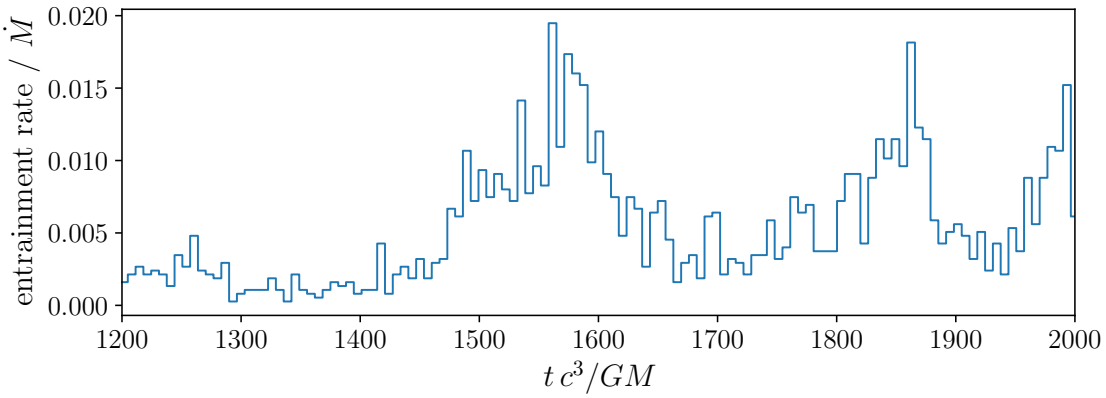


Figure 5.10: Histogram showing when tracer particles are entrained into the jet over a brief interval in the MAD $a_* = -0.94$ model. Entrainment is conservatively defined to only include particles that begin in the disk region and end at large radius with positive v^r. This definition discounts particles that spend time in the mixing region but ultimately fall onto the hole. In this MAD model and by these criteria, entrainment is evidently a stochastic process that is characterized by periods of increased entrainment corresponding to times when instabilities form and break at horizon scales.

Figure 5.10 shows the computed mass entrainment rate over time. We find that entrainment events occur in bursts lasting $\sim 100\,GM/c^3$. Mass loading occurs at an average rate $\sim 10^{-2}\,\dot{M}$. Note that our definition produces a measurement that does not count mass that has been injected by the numerical floor prescription in the funnel: the tracer particles are initialized once, so the application of floors during the subsequent evolution does not increase the number of the tracer particles. We discard the beginning epoch of tracer data to avoid including the floors' effect on the transient tracer particle initial condition.

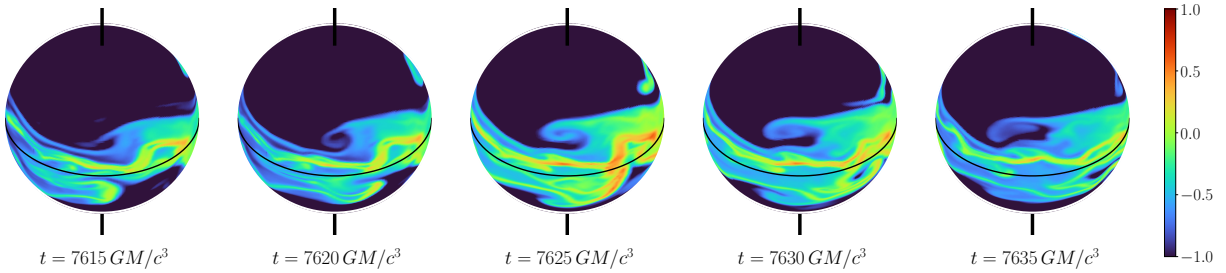


Figure 5.11: Logarithm over two decades of density on $r \approx 1.5\,M$ slices for the MAD $a_* = -0.94$ model at five times separated by $\Delta t = 25M$. Matter in the jet near the poles flows clockwise from above (left on the page), and matter in the midplane flows counterclockwise (right on the page). The boundary between the funnel and the midplane results in the development of an unstable shear layer. A Kelvin–Helmholtz roll develops in the shear layer over the sequence of panels.

In both SANE and MAD models, mass entrainment is driven by instabilities in the boundary between the accreting plasma and the matter in the jet. Figure 5.11 plots log plasma density on shells of constant radius over time and shows the development of an instability: as the high-density midplane disk region moves to the right, it interacts with the low density funnel plasma moving to the left and forms Kelvin–Helmholtz rolls. Figure 5.12 plots density and specific angular momentum in the central frame of Figure 5.11 in the $\theta - \phi$ plane at three different radii. Evidently, the KH roll is well resolved.

We observe that Kelvin–Helmholtz rolls develop in all simulations regardless of the accretion flow parameters; however, it is especially apparent in the MAD flows, which have a more turbulent boundary layer. Mass entrainment thus proceeds in part through the Kelvin–Helmholtz instability at the jet–disk boundary. Still, the full structure of the jet–disk boundary layer is complicated, and braked accretion streams near the event horizon also contribute to mass loading.

We also use the tracer particles to visualize the flow of matter through phase space. Figure 5.13 shows the time-averaged flow of tracer particles in the radius vs. specific angular momentum plane. Plasma density is represented by the density and thickness of the white flow lines. Color denotes particle speed in phase space and helps differentiate between the disk/plunging region and the jet.

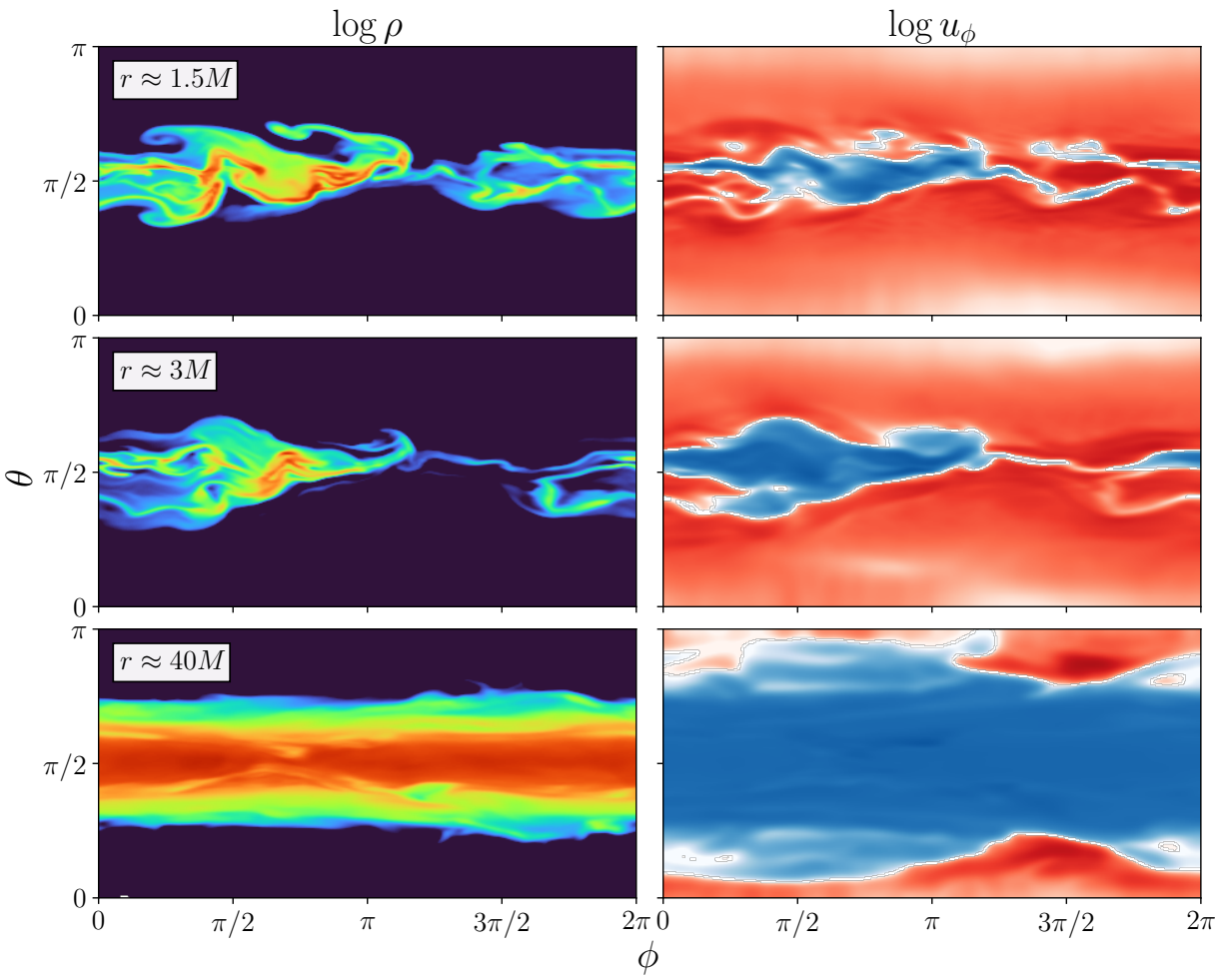


Figure 5.12: Left panels: log over two decades of density in the $\theta - \phi$ plane for shells at $r = 1.5, 3$, and $40\,GM/c^2$. Right panels: same shells as left showing logarithm over two decades of u_ϕ with $u_\phi > 0$ blue and red otherwise. These plots are from the central time slice of Figure 5.11, for the MAD $a_* = -0.94$ model. The flow becomes increasingly chaotic at smaller radii; however, the shear layer between the disk and funnel persists, and the funnel region consistently has $u_\phi < 0$, indicating corotation with the hole.

The flow at $r < 20$ can be divided into the three triangular regions shown in Figure 5.13. Region A contains particles that are falling toward the event horizon and gradually losing angular momentum. It contains the plunging region (where the figure is brightest), the disk, and the characteristic MAD accretion streams seen in Figure 5.4. Region B is the disk wind. Region C is the jet. Particles enter the jet from Region A, are torqued until their angular momentum has the same sign as the black hole, and then are accelerated outward. Particles

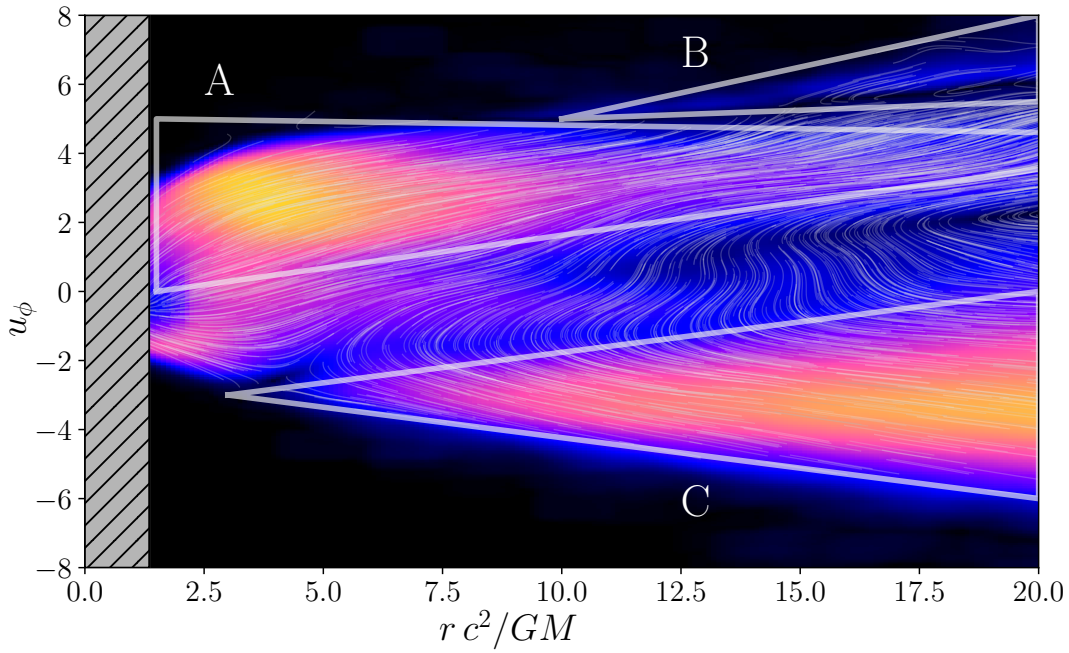


Figure 5.13: Time-averaged flow of tracer particles through the $r - u_\phi$ state space. The gray hatched region at the left of the figure lies within the horizon. The background shows a false-color representation of the average speed of the particles through the two-dimensional state space and helps to visually differentiate the disk (region A), disk wind (region B), and jet (region C). The density of white lines is proportional to the density of particles in state space; for the purposes of visualization, the density is capped for regions in the disk that have large density. Average particle flow follows the thin white lines. As particles are entrained in the jet they cross $u_\phi = 0$ and are then torqued and accelerate outwards.

gain angular momentum as they accelerate away from the hole, as expected in a sub-Alfvénic wind.

5.5 Discussion

We have studied a set of retrograde MAD and SANE black hole accretion models. We found that the angular momentum of plasma in both the jet and parts of the jet–disk boundary layer is aligned with the spin of the hole. We also found that the boundary layer region, in which u_ϕ transitions between its value in the midplane and its value in the jet, was wider in the MAD models than in SANE models. This is unsurprising, since MAD flows tend to be more chaotic near the horizon where much of the jet–disk interaction occurs, so the time-

averaged boundary location is spread out. The existence of a shear layer is not restricted to retrograde models, as noted in §5.2.2, but we have focused on retrograde models because the shear is strongest there.

As noted in §5.4.4, the jet–disk boundary is sufficiently resolved to see the development of Kelvin–Helmholtz rolls; this strongly suggests that numerical diffusion does not control the entrainment rate. Nevertheless increasing the simulation resolution may expose new structures, such as the plasmoids seen in recent high-resolution axisymmetric models (Nathanail et al., 2020; Ripperda et al., 2020).

To assess the effect of resolution we studied five different realizations of a MAD $a_* = -0.94$ model at four resolutions: one at 192 radial zones, one at 288, two at 384, and one at 448 (resolution in other coordinates is scaled proportionately). We include multiple realizations at the same resolution to assess the error bars on measurements associated with turbulent fluctuations. We consider convergence in two time-averaged quantities: the profiles of u_ϕ presented in Figure 5.9 and the total mass in the jet near the hole as measured from the GRMHD.

The time-averaged specific angular momentum profile $\langle u_\phi \rangle (r, \theta)$ is remarkably consistent across all resolutions everywhere except in the zones adjacent to the polar boundary, where we do not necessarily expect agreement because of our treatment of the boundary condition. In the shear region, the profiles are consistent to 5% and exhibit no discernible trend with resolution.

We compute the total mass in the jet near the hole by integrating the GRMHD density variable within a volume V

$$M_{\rm j}(t) \equiv \int_V \rho \sqrt{-g}\, {\rm d}r\, {\rm d}\theta\, {\rm d}\phi, \tag{5.20}$$

where we have chosen V to be the region with $u_\phi < 0$ and $v^r > 0$ at $2 < r < r_* = 20$. Note that $M_{\rm j}(t)$ has contributions from both mass entrainment and numerical floors. The

time-dependent variation in the entrainment rate (see Figure 5.10), causes $M_{\rm j}(t)$ to fluctuate, so evaluations of the time-averaged $\langle M_{\rm j}(t)\rangle_t$ are subject to noise. We find that $M_{\rm j}(t)$ has a correlation time $\approx 200\,GM/c^3$ in the MAD, $a_* = -0.94$ model. The full model duration is $20,000\,GM/c^3$, but the first $5,000\,GM/c^3$ is an unequilibrated transient, so we have $N \sim 80$ independent samples over the full model; therefore, we expect fractional errors of order $N^{-1/2} \sim 10\%$. We find that $\langle M_{\rm j}(t)\rangle = 140, 130, 160$, and 130 for simulations with radial resolution 192, 288, 384, and 448 respectively, which is consistent with the expected error. We also note that the widths of the jet and boundary-layer regions (in zones) reported in §5.4.3 scales linearly with the simulation resolution.

There may be additional mixing processes that occur on unresolved scales, so the consistency of $M_{\rm j}$ across resolutions does not prove that we have accurately accounted for mass mixing between the jet and disk. Future convergence studies should probe not only longer timescales to reduce the fluctuation noise but also higher resolution.

We also note that since the equilibration time increases with radius, the long-term average jet–disk interaction may be poorly represented at large radii where the disk is still strongly dependent on initial conditions. We have chosen to overstep this issue by only reporting fits and statistics from equilibrated parts of our simulations. Chatterjee et al. (2019) also studied mass loading in their study of black hole jet launching. They performed multiple long-time, large-scale ($r_{\rm max} \gtrsim 10^5\ GM/c^2$) 2D GRMHD simulations and found that additional mass entrainment occurred at large radii. As noted above, the details of the jet–disk interaction at such large radii may be influenced by the choice of initial condition.

5.6 Conclusion

We have studied a set of three-dimensional GRMHD simulations of retrograde SANE and MAD black hole accretion disks at $a_* = -0.5$ and -0.94, with a focus on the jet–disk boundary near the horizon. We have found that:

1. Plasma in the jet rotates with the hole and not the disk. This generates a jet–disk boundary with strong currents and vorticity.

2. In MAD models, accretion occurs through narrow plasma streams near the horizon. These streams erode as they interact with the counterrotating jet and load the jet with plasma.

3. In both MAD and SANE models, disk plasma is entrained into the jet in well-resolved Kelvin–Helmholtz rolls.

4. The entrainment rate is $\sim 0.01\,\dot{M}$ for the MAD, $a_* = -0.94$ model that we study in detail.

5. The entrainment rate and boundary-layer structure are insensitive to resolution over the range in resolution we are able to study.

6. In retrograde MAD models accretion near the horizon fluctuates strongly: individual time slices do not look like time- and azimuth-averaged data. Relatedly, the jet in MAD models wobbles significantly. The fluctuations produce a complicated, time-variable interface between the jet and disk.

This study has considered a limited range of models and could be extended by comparing a broader range of black hole spins and tilts between the hole and the accretion flow. Understanding the behavior of jet plasma and the jet–disk boundary layer may be crucial in developing a robust model of the connection between black hole spin and motion in the jet, which can now be resolved in time and space by the EHT.

Chapter 6

Drizzle pair production in black hole jets

Electron–positron pair creation near sub-Eddington accretion rate black holes is believed to be dominated by the Breit–Wheeler process (photon–photon collisions). The interacting high energy photons are produced when unscreened electric fields accelerate leptons either in coherent, macroscopic gaps or in incoherent structures embedded in the turbulent plasma flow. The latter type of acceleration results in a drizzle of pair production sourced by photons from the background radiation field whose energies are near the pair-production threshold. In this chapter, radiation GRMHD simulations are used to extend an earlier study of *pair drizzle* by Mościbrodzka et al. (2011). This study focuses on low-magnetization (SANE) accretion onto supermassive Kerr black holes and considers radiation due to synchrotron, bremsstrahlung, and Compton upscattering. In these simulations, pair drizzle in M87 is sufficient to keep the magnetospheric charge density orders of magnitude above the Goldreich–Julian density, and pair production peaks along the jet–disk boundary.

6.1 Introduction

The Event Horizon Telescope (EHT) recently published the first resolved images of plasma surrounding the M87 black hole at 1.3mm (Event Horizon Telescope Collaboration et al., 2019a,d). Although these images and anticipated future results carry information about physical conditions in the accreting plasma, an accurate model of the emission source—the radiating leptons—must be obtained in order to extract the information (see Event Horizon

This chapter is lightly adapted from G. N. Wong, B. R. Ryan, and C. F. Gammie, ApJ, Volume 907, Issue 2, id.73.

Telescope Collaboration et al., 2019e). The subset of the radiating leptons that originates as electron–positron pairs is of particular interest (see Svensson & Zdziarski 1989 for a review), and with this goal in mind, we consider a nearly *ab initio* model of pair production in low accretion rate (highly sub-Eddington) systems like M87 (see Broderick & Tchekhovskoy, 2015; Hirotani, 1989).

Models of pair production around low accretion rate black holes require a population of high-energy leptons that can Compton upscatter the low frequency background photons produced by the hot plasma. The upscattered high-energy photons produce electron–positron pairs through interactions with the fiducial low-energy background photons via the Breit & Wheeler (1934) process when the center-of-momentum energy of the interacting photons exceeds the rest-mass energy of an electron–positron pair ~ 1 MeV.

High-energy leptons can be produced in a variety of ways. Gap models envisage coherent regions with $\mathbf{E} \cdot \mathbf{B} \neq 0$ that accelerate the leptons and initiate pair cascades (see, e.g., Beskin et al. 1992; Hirotani & Okamoto 1998; Ford et al. 2018; Levinson & Cerutti 2018; Chen et al. 2018; Parfrey et al. 2019; Chen & Yuan 2020). In gap models, the high-energy photons typically have energies that are orders of magnitude above the MeV threshold. In contrast, *drizzle* models predict that the native high-energy component of the electron distribution throughout the near-horizon plasma will produce a steady, smooth background of $\sim$ MeV photons that interact with each other and pair produce (see Mościbrodzka et al. 2011, hereafter M11, and also Levinson & Cerutti 2018). Although gap and drizzle models may appear distinct, they can be thought of as end members of a continuum in which the structures that accelerate the leptons range from coherent, steady, and large scale (gap) to incoherent, transient, and small scale (drizzle).

In this paper we revisit the drizzle model of M11, which estimated pair production rates based on nonradiative general relativistic magnetohydrodynamics (GRMHD) accretion simulations. We extend the M11 estimate using *radiative* GRMHD (radGRMHD) accretion simulations produced with the `ebhlight` code (Ryan et al., 2015, 2019). `ebhlight` inde-

pendently tracks the ion and electron temperatures (Ressler et al., 2015; Ryan et al., 2017) and implements a more thorough treatment of electron thermodynamics that explicitly includes both a model to partition dissipation between electrons and ions and a treatment of ion–electron energy exchange through Coulomb scattering (Ressler et al., 2015). Moreover, `ebhlight` accurately accounts for radiative cooling by solving the radiation transport equation with a Monte Carlo method, which can be important to the plasma dynamics as accretion rates increase. Furthermore, in contrast to the M11 model, our pair production calculation includes photons produced by bremsstrahlung emission, which are unimportant for the thermal evolution of the fluid at the accretion rates we consider but may play an important role in drizzle pair production due to their characteristic high frequencies (Yarza et al., 2020). These extensions improve the accuracy of the pair production rate evaluation, especially at high accretion rates.

It is computationally expensive both to produce radGRMHD simulations and to generate well-resolved samples of the radiation field, so we evaluate pair drizzle for a targeted set of axisymmetric models. We consider two black hole spins $a_* \equiv Jc/GM^2 = 0.5$ and 0.94 (here J and M are the angular momentum and mass of the black hole, respectively) over a range of mass accretion rates $\dot{m} \equiv \dot{M}/\dot{M}_{\rm Edd}$[1] corresponding to geometrically thick, optically thin, weakly radiative accretion flows in the low magnetic flux "standard and normal evolution" (SANE) accretion state.[2]

The paper is organized as follows. Section 2 reviews the governing equations of radGRMHD, and Section 3 describes the pair production model and its implementation. In Section 4, we discuss the expected spatial dependence of drizzle pair production. Section 5 presents the results of our numerical simulations. We discuss physical implications and model limitations in Section 6, and we provide a summary in Section 7.

[1]Here, $\dot{M}_{\rm Edd} \equiv L_{\rm Edd}/\left(\eta\, c^2\right)$ where $\eta = 0.1$ is the nominal accretion efficiency, and the Eddington luminosity $L_{\rm Edd} \equiv 4\pi GMm_p c/\sigma_T$

[2]SANE in contrast to "magnetically arrested disk" (MAD) models, which have magnetic flux through the event horizon Φ satisfying $\phi \equiv \Phi/(\dot{M}r_g^2 c)^{1/2} \simeq 15$; here $r_g = GM/c^2$.

6.2 Plasma model

We consider prograde black hole accretion, in which the orbital angular momentum of the plasma is aligned with the spin of the central black hole. Our models have accretion rates $\dot{m} \leq 10^{-5}$. We find that these accretion rates are low enough for the plasma to be $\sim$ collisionless (i.e., the Coulomb scattering mean free path for electrons and ions is large compared to GM/c^2) but high enough that radiative cooling may influence the electron temperature. Hereafter, we set $GM = c = m_e = 1$ and occasionally restore cgs units for clarity. We model the plasma using radiative general relativistic radGRMHD.

In a coordinate basis, the governing equations of radGRMHD are

$$\partial_t \left(\sqrt{-g}\rho_0 u^t\right) = -\partial_i \left(\sqrt{-g}\rho_0 u^i\right), \tag{6.1}$$

$$\begin{aligned}\partial_t \left(\sqrt{-g}T^t{}_\nu\right) = \partial_i \left(\sqrt{-g}T^i{}_\nu\right) + \sqrt{-g}T^\kappa{}_\lambda \Gamma^\lambda{}_{\nu\kappa} \\ - \sqrt{-g}R^\mu{}_{\nu;\mu},\end{aligned} \tag{6.2}$$

$$\partial_t \left(\sqrt{-g}B^i\right) = \partial_j \left[\sqrt{-g}\left(b^j u^i - b^i u^j\right)\right], \tag{6.3}$$

$$\partial_i \left(\sqrt{-g}B^i\right) = 0, \tag{6.4}$$

where the plasma is defined by its rest-mass density ρ_0, its four velocity u^μ, and b^μ is the magnetic field four vector following McKinney & Gammie (2004). Here, $g \equiv \det(g_{\mu\nu})$ is the determinant of the covariant metric, Γ is a Christoffel symbol, and i and j denote spatial coordinates. In Equations 6.3 and 6.4, we express components of the electromagnetic field tensor $F^{\mu\nu}$ as $B^i \equiv \star F^{it}$ for notational simplicity.

The stress-energy tensor $T^\mu{}_\nu$ contains contributions from both the fluid and the electromagnetic field:

$$\begin{aligned}T^\mu_\nu = \left(\rho_0 + u + P + b^\lambda b_\lambda\right) u^\mu u_\nu \\ + \left(P + \frac{b^\lambda b_\lambda}{2}\right) g^\mu_\nu - b^\mu b_\nu,\end{aligned} \tag{6.5}$$

where u is the internal energy of the fluid and the fluid pressure P is related to its internal energy through an adiabatic index $\hat{\gamma}$ with $P \equiv (\hat{\gamma} - 1)\, u$ (see, e.g., Gammie et al. 2003; McKinney & Gammie 2004).

The radiation stress tensor is

$$R^{\alpha}{}_{\beta} = \int \frac{d^3p}{\sqrt{-g}p^t} p^{\alpha} p_{\beta} \left(\frac{I_{\nu}}{h^4 \nu^3} \right), \tag{6.6}$$

where p^{α} is the four-momentum of a photon, ν is the frequency of the photon, I_{ν} is specific intensity, and h is Planck's constant. Photons obey the equations of radiative transfer equations as they move through the plasma:

$$\frac{dx^{\alpha}}{d\lambda} = k^{\alpha}, \tag{6.7}$$

$$\frac{dk^{\alpha}}{d\lambda} = -\Gamma^{\lambda}{}_{\alpha\beta} k^{\alpha} k^{\beta}, \tag{6.8}$$

$$\frac{D}{d\lambda} \left(\frac{I_{\nu}}{\nu^3} \right) = \frac{\eta_{\nu}(T_e)}{\nu^2} - \frac{I_{\nu} \chi_{\nu}(T_e)}{\nu^2}. \tag{6.9}$$

Here, η_{ν} is the local emissivity of the plasma, and χ_{ν} encodes the total (scattering and absorption) opacity due to thermal synchrotron processes and Compton scattering.

We consider a two-temperature plasma composed of electrons and ions. The extra degree of freedom is closed through an independent electron energy equation as in Ressler et al. (2015):

$$\frac{\rho^{\hat{\gamma}_e}}{\hat{\gamma}_e - 1} u^{\mu} \partial_{\mu} \kappa_e = f_e Q_H + Q_C(T_e, T_p) - u^{\nu} R^{\mu}{}_{\nu;\mu}, \tag{6.10}$$

where $\hat{\gamma}_e$ is the adiabatic index of the electrons, f_e is the fraction of the volumetric dissipation rate Q_H that goes into electrons (taken from the Howes 2010 model), and Q_C is volumetric heating or cooling due to Coulomb scattering (Stepney & Guilbert, 1983).

6.2.1 Two-temperature radGRMHD

We solve the governing equations using the `ehblight` code (Ryan et al., 2017). In `ebhlight`, the radiation stress-energy tensor is co-evolved with the fluid and is computed at each step from Monte Carlo samples of the radiation field, which are evolved according to the scheme introduced in `grmonty` (Dolence et al., 2009). `ebhlight` also independently tracks the proton and electron temperatures according to a two-temperature model where the electron entropy is evolved as in Ressler et al. (2015).

6.3 Pair production

Pair drizzle in low accretion rate systems is weak, so the radiation field can be treated as independent of pair production. The pair production rate can thus be evaluated in a post-processing step after the fluid evolution has been completed. We will show below that this approximation is self-consistent.

6.3.1 Comparison of contributing interactions

The pair production rate density is

$$\dot{n}_{\pm} = n_1 \, n_2 \, \langle \sigma_{12} \, v \rangle \tag{6.11}$$

where n_1 and n_2 are the number densities of the two interacting species, σ_{12} is their interaction cross section, v is their relative velocity, and the angle brackets indicate an average over state variables.

In our radiative electron–ion plasma, pair-producing interactions can occur between electrons (e), ions (p), and photons (γ). The pair production cross sections are (Phinney, 1983;

Stepney & Guilbert, 1983; Zdziarski, 1985; Krolik, 1999)

$$\sigma_{pp} \sim \sigma_{ee} \sim \sigma_{ep} \sim \alpha\sigma_{p\gamma} \sim \alpha\sigma_{e\gamma} \sim \alpha^2\sigma_{\gamma\gamma}, \tag{6.12}$$

where $\alpha \approx 1/137$ is the fine-structure constant. Which process dominates depends on the details of the radiation field and plasma density. We can estimate n_e using `ebhlight` simulations for guidance and assuming a pure hydrogen plasma; we can also estimate n_γ by analyzing the simulated radiation field and counting only photons with energy $> m_e c^2$. Then in the low density *jet* region near the spin axis of the black hole, we find that $n_\gamma/n_e = n_\gamma/n_p > 1 > \alpha$.[3] The $\gamma\gamma$ process therefore dominates pair production.

6.3.2 Basic equations

The pair production rate density due to the $\gamma\gamma$ process (counting pairs and not individual particles) is

$$\begin{aligned}\dot{n}_\pm \equiv& \frac{1}{\sqrt{-g}} \frac{\mathrm{d}N_\pm}{\mathrm{d}^3x\,\mathrm{d}t} \\ =& \frac{1}{2} \int \frac{\mathrm{d}^3k}{\sqrt{-g}} \frac{\mathrm{d}^3k'}{\sqrt{-g}} \frac{\mathrm{d}N_\gamma}{\mathrm{d}^3x\,\mathrm{d}^3k} \frac{\mathrm{d}N_\gamma}{\mathrm{d}^3x\,\mathrm{d}^3k'} \frac{\epsilon^2}{k^0 k'^0}\, \sigma_{\gamma\gamma}\, c,\end{aligned} \tag{6.13}$$

where $\mathrm{d}N_\gamma/\mathrm{d}^3x\,\mathrm{d}^3k$ is the photon distribution function, the factor of $1/2$ prevents double counting of interacting photons, and the center-of-momentum energy ϵ and cross section $\sigma_{\gamma\gamma}$ are

$$\epsilon^2 = -\frac{k_\mu k'^\mu}{2}, \tag{6.14}$$

$$\begin{aligned}\sigma_{\gamma\gamma} = \frac{3\sigma_T}{8\epsilon^6} \Big[&\left(2\epsilon^4 + 2\epsilon^2 - 1\right) \cosh^{-1}\epsilon - \\ &\left(\epsilon^3 + \epsilon\right) \sqrt{\epsilon^2 - 1} \Big],\end{aligned} \tag{6.15}$$

[3] `ebhlight` can resolve only a limited density contrast, so the density is artificially increased in the jet via numerical "floors." The numerical electron density is therefore an upper limit on the physical density.

where σ_T is the Thomson cross section (see Breit & Wheeler, 1934). Note that the phase space volume element $\mathrm{d}^3x\,\mathrm{d}^3k$ is only invariant if the integration is over the components of the covariant wave four vector, i.e., $\mathrm{d}^3k \equiv \mathrm{d}k_1\,\mathrm{d}k_2\,\mathrm{d}k_3$.

The photon distribution function (i.e., the radiation field) is generated by the same synchrotron emission and absorption plus Compton scattering physics of the radGRMHD model; however, for the pair computation we include an additional model for bremsstrahlung emission (bremsstrahlung absorption is negligible). We note that bremsstrahlung is energetically subdominant everywhere in our models, but it may be an important source of high-energy (and therefore pair-producing) photons. We adopt the piecewise bremsstrahlung emissivity of Straub et al. (2012) (see also Yarza et al., 2020).

6.3.3 Numerical implementation of pair production

In radGRMHD, the plasma evolution depends on the radiation stress-energy tensor, which is an integral over the entire photon distribution function. In contrast, the $\gamma\gamma$ pair production rate is a double integral over the photon distribution function and is dominated by a small range of energies around the pair production threshold. The pair production rate calculation therefore requires a more accurate estimate of the photon distribution function than does the plasma evolution. This is the main numerical motivation for evaluating the pair production in post-processing.

Our procedure is as follows. We generate a detailed sample of the radiation field using a Monte Carlo step that re-simulates the radiative transport and includes bremsstrahlung emission.[4] In this scheme, each radiation field sample i is assigned a weight w_i equal to the number of physical photons in the sample multiplied by a constant that is inversely proportional to photon sampling cadence. Given a list of radiation field samples within a

[4]We use a "fast-light" approximation, which neglects the light-crossing time and allows us to avoid coupling snapshots of the plasma state taken at different coordinate times.

cell of coordinate volume $\Delta^3 x$,

$$\dot{n}_{\pm} \approx \frac{1}{2} \sum_{i,j} \frac{w_i}{\sqrt{-g}\Delta^3 x} \frac{w_j}{\sqrt{-g}\Delta^3 x} \frac{\epsilon^2}{k_i^0 k_j^0} \sigma_{\gamma\gamma}\, c, \tag{6.16}$$

where the Latin indices label radiation field samples in that cell.

Although Equation (6.16) can be summed pairwise over all n samples, it is more efficient to sample the sum over a subset of $m < n^2$ pairs (i, j). We set an upper limit on the number of pairs to consider and use reservoir sampling to obtain an unbiased subset from the full list. In the limit that the m is large, the error in Equation (6.16) exhibits the usual $m^{-1/2}$ Monte Carlo scaling.

6.3.4 Test Problems

We now consider two tests to verify our method. Both tests comprise two steady, isotropic, pointlike photon sources in flat space separated by a distance $2L$. For each case, we measure the pair production rate along the perpendicular bisector of the line connecting the two sources (see Figure 6.1).

Monochromatic point sources

In the first test, each point source is monochromatic, and the pair production rate can be evaluated analytically (see Mościbrodzka et al. 2011; the test is provided here as a consistency check). The center of mass energy ϵ is

$$\epsilon^2 = -\frac{k_\mu k'^\mu}{2} = \frac{k^0 k'^0 \left(1 - \cos\theta\right)}{2}, \tag{6.17}$$

and

$$\frac{\mathrm{d}N_\gamma}{\mathrm{d}^3 x} = \frac{\dot{N}_\gamma \sin^2\left(\theta/2\right)}{4\pi L^2 c}, \tag{6.18}$$

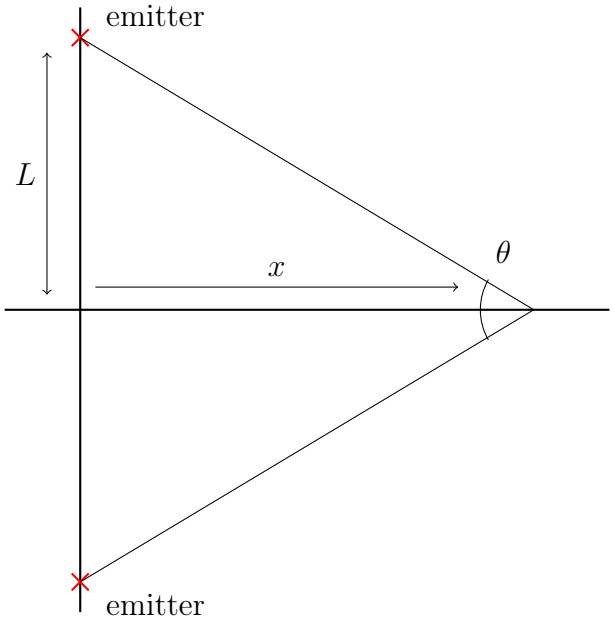


Figure 6.1: The test problem geometry comprises two isotropic emitters separated by a distance $2L$. The pair production rate density $\dot{n}_{\pm}(x)$ is evaluated as a function of distance x along the perpendicular bisector of the two sources. The angle between two incident photons at a point x along the bisector is $\theta = 2\arctan(L/x) \approx 2L/x$ for $x/L \gg 1$.

where $\dot{N}_{\gamma}$ is the rate of isotropic photon production in the frame of the emitters. Then

$$\dot{n}_{\pm}(\theta) = \left(\frac{\dot{N}_{\gamma} \sin^2 \theta}{4\pi L^2 c} \right)^2 (1 - \cos\theta)\, \sigma_{\gamma\gamma}(\epsilon)\, c. \tag{6.19}$$

Figure 6.2 compares $\dot{n}_{\pm}(x)$ computed analytic versus numerically in the upper panel and shows the fractional difference between the two evaluations in the lower panel. Figure 6.3 shows the fractional difference between the domain-averaged numerical values and the analytic expression as a function of N_s, the number of samples of the radiation field. As expected, the error scales as $N_s^{-1/2}$.

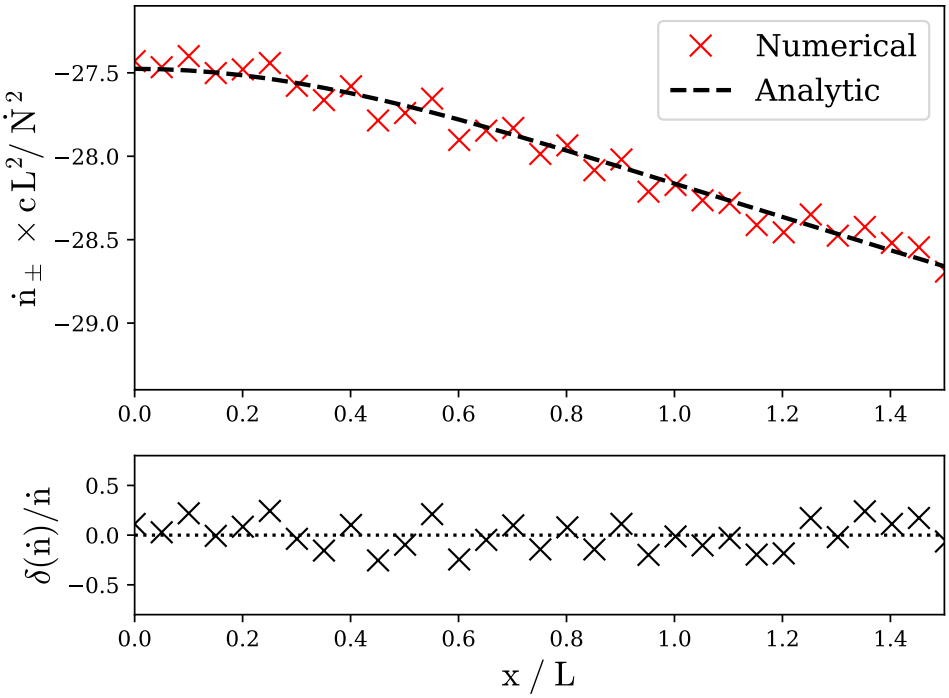


Figure 6.2: Monochromatic emitter test problem for $N_s \approx 10^6$. Upper panel: numerical (red hashes) and analytic (black line) pair production rate densities for two monochromatic, isotropic emitters with source separation 2L, evaluated as a function of radius in the plane normal to and bisecting the line connecting the emitters. Lower panel: fractional difference between numerical and analytic values.

Power law spectrum point sources

In the second problem, we endow each point source with a power law spectrum with index α and cutoff frequencies $\nu_{\rm min} \ll \nu_e$ and $\nu_{\rm max} \gg \nu_e$, where $\nu_e \equiv m_e c^2/h$. In particular,

$$L_\nu = \begin{cases} \frac{L_0}{\nu_e}\left(\frac{\nu}{\nu_e}\right)^\alpha & \nu_{\rm min} < \nu < \nu_{\rm max} \\ 0 & \text{otherwise.} \end{cases} \tag{6.20}$$

The pair production cross section peaks for $\epsilon \sim 1$, and so the dominant contribution from photons with $\nu_{\rm min} \ll \nu_e$ will be through their interactions with high energy photons at frequencies $\nu = \nu_e/\nu_{\rm min}$, provided α is not too large (otherwise most pair production is by photons with $\nu \sim \nu_{max}$). In the astrophysical settings of interest to us, $\nu_e/\nu_{\rm min} \gg \nu_{\rm max}/\nu_e$,

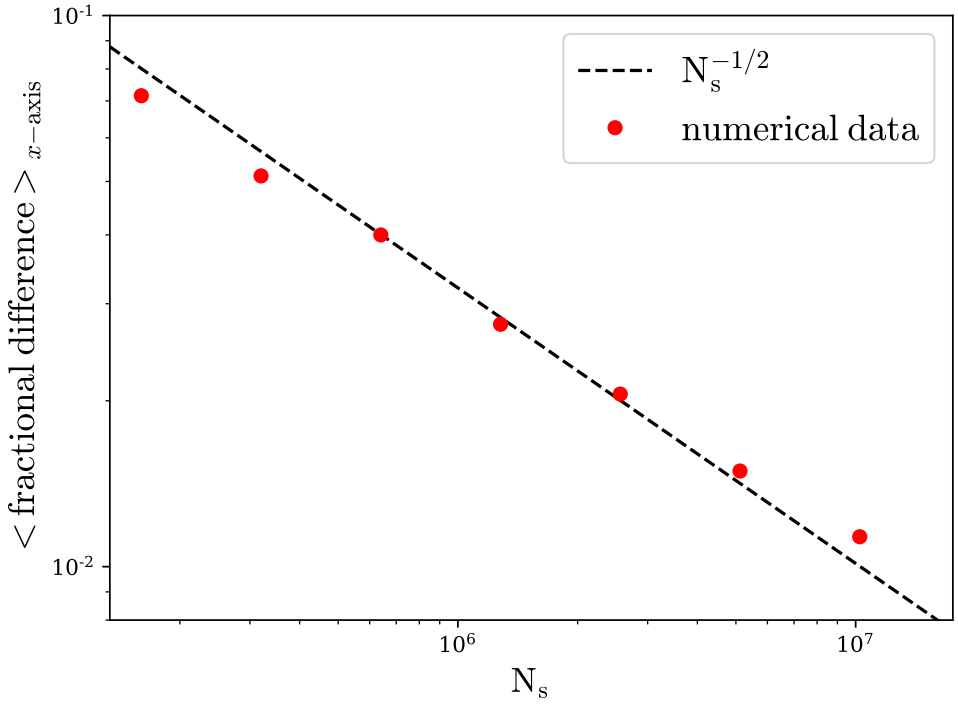


Figure 6.3: Code convergence. Averaged fractional difference between numerical and analytic pair production rate densities for the two-point, monochromatic, isotropic emitter problem as a function of number of field samples generated. The error scales $\propto N_s^{1/2}$ as expected.

so we can neglect the dependence of $\dot{n}_\pm$ on $\nu_{\min}$.

Analytic evaluation of the pair production rate density for this test is difficult because the pair production cross section depends on energy in a nontrivial way, but the asymptotic scaling with x is easy to compute. At each $\theta = 2\arctan(L/x) \approx 2L/x$ for $x \gg L$, the dominant contribution to the rate integral is at $\epsilon \sim 1$ or $\nu \sim (\nu_e^2/\nu')(4x^2/(L^2))$. The product of the distribution functions thus scales as $x^{-4+2\alpha}$ because of the relationship between ν and ν' and $k^0 k'^0 \sim \nu\nu' \sim x^2$, implying[5] that

$$\dot{n}_\pm \sim x^{-6+2\alpha}. \tag{6.21}$$

The radial dependence of the pair production rate is therefore a nontrivial function of the

[5]Rigorously: take the cross section to be a δ function in ϵ and integrate over ν.

source spectral index measured at pair-producing energies. Figure 6.4 compares the numerically evaluated pair production rate density to the analytic estimate.

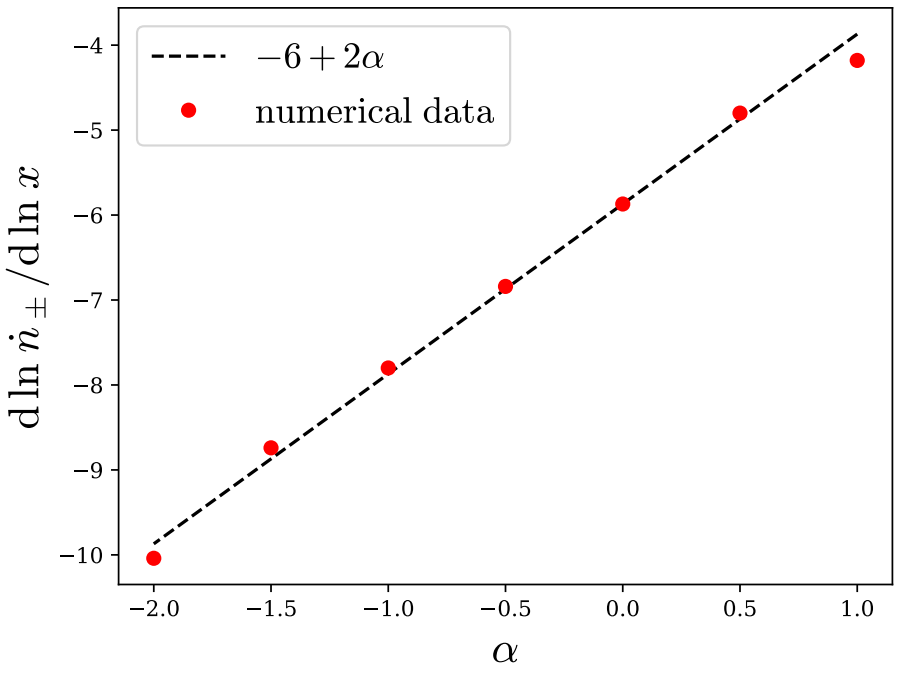


Figure 6.4: Radial dependence of pair production rate density vs. source spectrum index. Here, α is the index of the source radiation spectrum $L_\nu \sim \nu^\alpha$, and the slope $\mathrm{d}\ln \dot{n}_\pm/\mathrm{d}\ln x$ describes the asymptotic radial power law dependence of the pair production rate density versus distance x from the source. The numerical results are plotted against and agree with the analytic estimate.

6.3.5 Goldreich–Julian charge density

There may be regions in a black hole magnetosphere where the charge density is insufficient to screen electric fields in the frame of the plasma and thus where the ideal MHD condition is violated. In these regions, the unscreened electric field can accelerate electrons and positrons to sufficiently high energies that they produce photons above the pair production threshold. Once this new generation of electron–positron pairs is produced, they themselves are accelerated in the unscreened electric field. This process can repeat over multiple generations, and

ultimately, the *pair cascade* will continue until enough charge has been produced to short out the potential (see, e.g., Sturrock 1971; Ruderman & Sutherland 1975 and Beskin et al. 1992 in the context of black holes).

The minimum charge density required to screen electric fields is known as the Goldreich–Julian charge density $n_{\rm GJ}$ (Goldreich & Julian, 1969). In covariant language, a charge density ρ_q is given by $\rho_q = -u^\mu j_\mu$, where j^μ is the four current and u^μ is the four velocity of the frame in which the charge density is measured; ρ_q is thus a frame-dependent quantity. The four current is *always* given by Maxwell's equations $j^\mu = F^{\mu\nu}{}_{;\nu}$ (notice that the covariant derivative includes time derivatives).[6] Goldreich and Julian's calculation is done in flat space, and the charge density is measured in the nonrotating frame. In our case, the choice of frame is less obvious. If the magnetosphere solution can be described by ideal MHD, then the uniquely sensible choice of frame is the fluid frame. In the Blandford–Znajek (BZ; Blandford & Znajek 1977) solution, however, there is no unique four-velocity associated with the force-free solution. Instead, we evaluate the charge density—which we will call the Goldreich–Julian density—in the normal observer frame:

$$n_{\rm GJ} e = -n^\mu j_\mu. \tag{6.22}$$

In the normal observer frame $u_\mu = n_\mu \propto (1,0,0,0)$ and using j^μ as the four current of the BZ split monopole solution,

$$n_{\rm GJ} \approx \left(\frac{a_* B^r c^2}{4\pi G M e}\right) \frac{(1+2/x)^{1/2}\cos\theta}{x^3}, \tag{6.23}$$

where $x \equiv r/\mathcal{L}$, $\mathcal{L} \equiv GM/c^2$ and B^r is now the radial component of the magnetic field at $x = 1$ in spherical Kerr-Schild coordinates (see Mościbrodzka et al., 2011).

[6] When calculating j^μ from a GRMHD simulation, we first construct $F^{\mu\nu}$ from the four-velocity u^μ, the magnetic induction four-vector b^μ, and the ideal MHD condition $F^{\mu\nu}u_\nu = 0$ according to $F^{\mu\nu} = \epsilon^{\mu\nu\alpha\beta}u_\alpha b_\beta$, where $\epsilon^{\mu\nu\alpha\beta}$ is the Levi–Civita tensor. The four current can then be computed from the inhomogeneous Maxwell equations, $j^\mu = F^{\mu\nu}{}_{;\nu}$. We use a finite-difference method across neighboring simulation locations and time slices to evaluate the derivative.

Assuming that the magnetic pressure $\sim$ the gas pressure and that both are of order ρc^2, then for $B^r \approx 10^4\sqrt{\dot{m}/m_8}$, where $m_8 \equiv M/\left(10^8 M_\odot\right)$. The charge number density for $a_* = 0.94$ is

$$n_{\rm GJ} \simeq 2.0 \times 10^{-1}\, \dot{m}^{1/2}\, m_8^{-3/2}\, {\rm cm}^{-3}. \tag{6.24}$$

Notice that the charge density does not necessarily vanish in MHD, since $\mathbf{E} = 0$ does not imply $\nabla \cdot \mathbf{E} \neq 0$. Nevertheless, the MHD solution cannot be self-consistent where $n < n_{\rm GJ}\,(= -u^\mu j_\mu/e)$, since Maxwell's equations cannot be satisfied. Also note that if $n \gg n_{GJ}$, then it is not clear how to produce macroscopic regions with unscreened electric fields (gaps) in MHD unless there is an unresolved process that drives the number density toward zero. Finally, note that our ideal MHD simulations can never represent macroscopic regions where $\mathbf{E} \cdot \mathbf{B} \neq 0$, so they are incapable of recovering the dynamics of gaps and pair cascades.

6.4 Spatial distribution of pair production

We now provide a simple geometric treatment to motivate the function form of the drizzle pair production rate density due to each component of the background radiation spectrum. In the following section, we will use the numerical results to fit the model parameters. In general, the drizzle pair production rate density may be a function of time and space, and it can depend on model parameters like black hole spin. Because of symmetries in the spacetime, however, we expect the *mean* rate density to be independent of time and azimuth.

The density of pair-producing photons is

$$n_\gamma = n_{\gamma,{\rm synch}} + n_{\gamma,{\rm Compt}} + n_{\gamma,{\rm brems}}, \tag{6.25}$$

where the terms represent photons produced by direct synchrotron emission, Compton scattering, and direct bremsstrahlung emission respectively. In our models, scattered bremsstrahlung

photons and direct synchrotron photons are both negligible near the pair-production threshold.

Since $n_{\gamma,\mathrm{Compt}}$ and $n_{\gamma,\mathrm{brems}}$ have different spatial distributions, we neglect the $n_{\gamma,\mathrm{Compt}} n_{\gamma,\mathrm{brems}}$ cross term, which is negligible compared to $n^2_{\gamma,\mathrm{Compt}}$ and $n^2_{\gamma,\mathrm{brems}}$ over the bulk of the domain. We thus approximate the total drizzle pair production rate density as a sum of two independent terms due to self-interaction of Compton and bremsstrahlung photons, respectively,

$$\dot{n}_{\pm}(r,\mu) \approx \dot{n}_{\pm,\mathrm{Compt}} + \dot{n}_{\pm,\mathrm{brems}}. \tag{6.26}$$

6.4.1 Compton contribution

The pair production rate density is a strongly decreasing function of distance from the photon source. Compton upscattered pair-producing photons come from regions of high electron temperature, and so the Compton contribution is likely to correlate strongly with regions of peak electron temperature, which exist both in the jet–disk boundary layer (see the Wong et al. 2021a companion paper for a study of the jet–disk boundary layer) and close to the event horizon. Thus, the Compton contribution may be written as

$$\begin{aligned}\dot{n}_{\pm,\mathrm{Compt}}(r,\mu) = \mathcal{A}\left(\frac{r}{\mathcal{L}}\right)^{-\alpha}\Big(e^{-\mu^2/2\sigma^2} + \\ \mathcal{B}\left(e^{-\left(\mu-\mu_f\right)^2/2\sigma_f^2} + e^{-\left(\mu+\mu_f\right)^2/2\sigma_f^2}\right)\Big)\end{aligned} \tag{6.27}$$

where $\mathcal{A}$ is an overall normalization with dimensions of rate density, $\mathcal{B}$ describes the relative importance of the jet–disk boundary versus the midplane, $\mu \equiv \cos\theta$, σ and σ_f describe the scale heights of pair production in the disk and the boundary layer, respectively, and μ_f is the location of the boundary layer. Following M11, we parameterize

$${\mu_f}^2 = \frac{r+a}{r+b}. \tag{6.28}$$

This model has seven parameters, $\mathcal{A}, \alpha, \mathcal{B}, \sigma, \sigma_f, a$, and b; however, we will find that the last four parameters can be fixed.

6.4.2 Bremsstrahlung contribution

Bremsstrahlung photons near the pair-production threshold are emitted primarily in regions where the dimensionless electron temperature $\Theta_e \equiv k_{\rm B} T_e / m_e c^2 \gtrsim 1/2$ (here $k_{\rm B}$ is Boltzmann's constant). In our models, this region extends out to approximately $r = 10\, GM/c^2 \equiv r_{\rm crit}$ and corresponds physically to the domain in which viscous heating, electron cooling, and Coulomb cooling are in approximate balance. Because bremsstrahlung emissivity depends only on n_i, n_e, and Θ_e, the geometry of the bremsstrahlung-driven region of pair production varies little from model to model.

Foresight from the numerical simulations and the presence of a radial cutoff at $r_{\rm crit}$ suggest that

$$\dot{n}_{\pm,\rm brems}(r,\mu) = \mathcal{C}\,\frac{\zeta^2}{\zeta^2+\mu^2}\begin{cases} \left(\dfrac{r}{\mathcal{L}}\right)^{-\kappa_1} & r < r_{\rm crit} \\ \left(\dfrac{r}{\mathcal{L}}\right)^{-\kappa_2} r_{\rm crit}^{\kappa_2-\kappa_1} & r_{\rm crit} \le r, \end{cases} \tag{6.29}$$

where ζ parameterizes the dependence of the distribution on elevation μ. This model has five parameters, $\mathcal{C}, \kappa_1, \kappa_2, r_{crit}$, and ζ. We find that fixing $\zeta = 0.4$ globally does not affect the quality of the fit.

The characteristics of the fluid and radiation field determine κ_1. The bremsstrahlung spectrum just above its peak follows a power law according to the behavior of Θ_e in the domain of emission, with $\alpha \approx -1$. Using Equation (6.21), we therefore expect $\kappa_2 = 6{+}2{\times}1 = 8$.

Model	a_*	m_8	$\dot{m}$	$L_{\rm bol}/L_{\rm Edd}$	$\epsilon_{\rm rad}$	$L_{\pm}/\left(L_{\rm BZ}\Gamma_j\right)$	l_c	notes
A5	0.5	33	2.2×10^{-5}	4.7×10^{-6}	0.021	1.0×10^{-5}	0.1	∼M87
A9	0.94	33	8.2×10^{-6}	1.5×10^{-6}	0.018	3.9×10^{-6}	0.03	∼M87
B5	0.5	62	9.2×10^{-6}	7.1×10^{-7}	7.7×10^{-3}	6.0×10^{-7}	0.02	∼M87
B9	0.94	62	5.2×10^{-6}	5.6×10^{-7}	1.1×10^{-3}	6.7×10^{-7}	0.01	∼M87
C	0.5	1	1.1×10^{-5}	7.1×10^{-7}	6.5×10^{-3}	1.8×10^{-7}	0.02	—
D	0.5	1	1.0×10^{-6}	1.3×10^{-8}	1.3×10^{-3}	4.6×10^{-9}	3×10^{-4}	—
E	0.5	1	1.3×10^{-7}	2.9×10^{-10}	2.2×10^{-4}	2.7×10^{-15}	7×10^{-6}	—
F	0.5	1	1.2×10^{-8}	4.2×10^{-12}	3.5×10^{-5}	3.4×10^{-18}	1×10^{-7}	—

Table 6.1: Time-Averaged RadGRMHD Model Parameters. From left to right: model name, dimensionless black hole spin parameter a_*, $m_8 \equiv$ black hole mass in units of $10^8 M_\odot$, $\dot{m} \equiv$ black hole accretion rate in units of Eddington mass accretion rate $\dot{M}_{\rm Edd} = 2.22\, m_8\ M_\odot\,{\rm yr}^{-1}$, time-averaged ratio of bolometric luminosity to Eddington luminosity, time-averaged radiative efficiency $\epsilon_{\rm rad} = L_{\rm bol}\,\dot{M}^{-1}c^{-2}$, ratio of rest-mass pair luminosity to BZ luminosity, and compactness parameter (related to efficiency of pair production, see Section 6.6.4 and Equation 6.37). The values reported in this table include the bremsstrahlung contribution and thus differ from previous results.

6.5 Numerical results

We now describe the results of our numerical pair drizzle simulations and provide fits for the model parameters described in § 6.4.

6.5.1 Simulation parameters

We consider the eight SANE radGRMHD models listed in Table 6.1. The first four models have M87-like parameters (Ryan et al., 2018) with varying spin and mass. The second four models increase $\dot{m}$ at fixed mass and spin until radiative cooling becomes important (Ryan et al., 2017).

The initial conditions for the fluid were produced by axisymmetrizing three-dimensional nonradiative GRMHD models, which are less computationally expensive to evolve. In mapping from three to two dimensions, the no-monopoles constraint was enforced by computing B^i from an axisymmetrized vector potential calculated from the original GRMHD simulations.

All fluid calculations were carried out in the modified Kerr-Schild (MKS) coordinates of

McKinney & Gammie (2004) with $h = 0.3$. The inner boundary was located within the event horizon, and the outer boundary was set at $r = 200\,GM/c^2$. The simulations were run at a resolution of 388 radial zones by 256 elevation zones.

6.5.2 Simulation outcomes

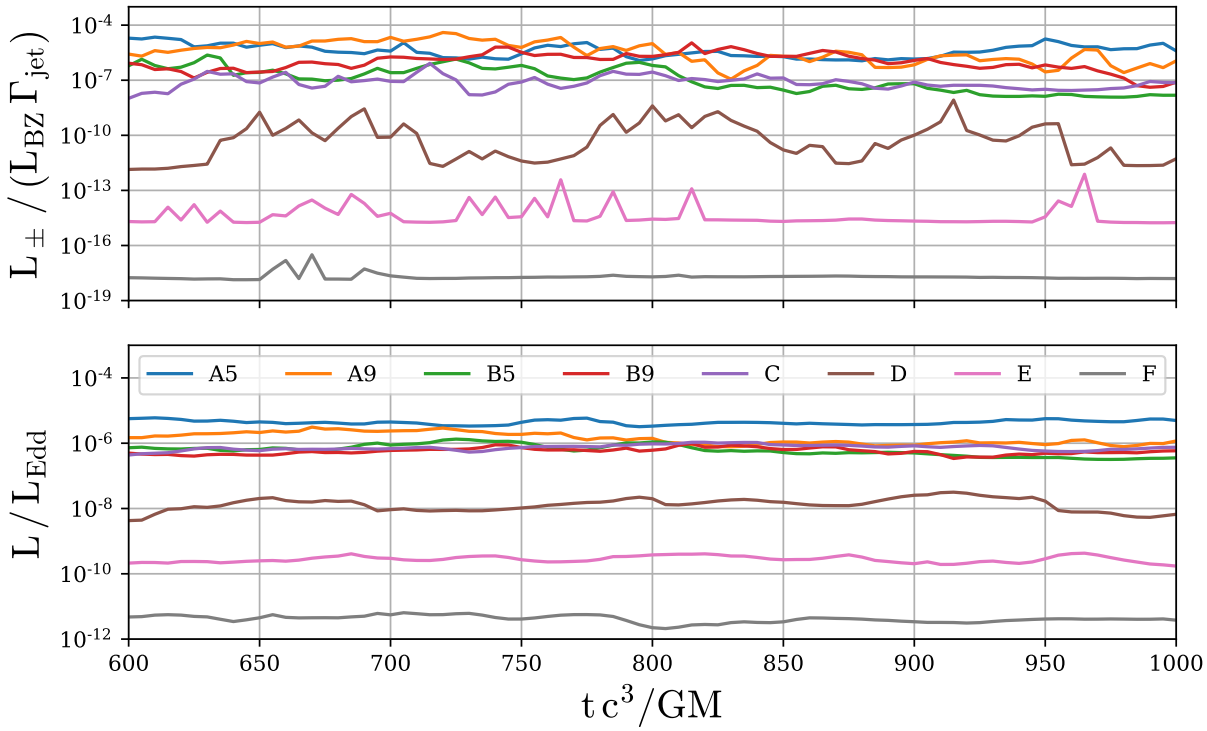


Figure 6.5: Time series of pair production rate and luminosity. Top: rest-mass pair drizzle luminosity divided by BZ jet power. Bottom: numerically calculated bolometric luminosity vs. time. Over our range of models, time variability increases with $\dot{m}$ because the increasingly important Compton contribution scales more favorably than the bremsstrahlung one.

In our models the cumulative pair production rate fluctuates over four orders of magnitude, with both the domain-integrated and position-dependent pair production rates varying on timescales as short as the fluid dump cadence $5\,M$. Following the discussion in Section 6.4.1, we find that the time- and azimuth-averaged $\dot{n}_{\pm,\mathrm{Compt}}$ peaks in the midplane and in *hotspot* regions (characterized by high Θ_e) that lie within the jet–disk boundary layer.

The structure and locations of the hotspot regions are highly variable.

The pair drizzle luminosity is

$$L_{\pm} \equiv 2 m_e c^2 \, \dot{N}_{\pm} \, \Gamma_{\rm jet}, \tag{6.30}$$

where $\Gamma_{\rm jet}$ is the bulk Lorentz factor at large r and the domain-integrated pair production rate is $\dot{N}_{\pm} \equiv \int \sqrt{-g} \, \mathrm{d}^3 x \, \dot{n}_{\pm}$. Figure 6.5 shows the time variability in both the pair drizzle luminosity and the background bolometric luminosity. Although both quantities exhibit variations, fluctuations in the former occur on shorter timescales and with greater amplitude. These variations are primarily caused by transient hotspot regions associated with plasmoids that form within the jet–disk boundary layer and travel across the domain. For models in which bremsstrahlung is the primary source of photons near the pair production threshold, variability is decreased. This is particularly evident in model F. For models in which Compton upscattering is the primary source of photons near the pair production threshold, time variability decreases as $\dot{N}_{\pm}$ increases. This is unsurprising, since increasing the pair production rate requires a larger fraction of the domain to be in a steadily pair-producing regime.

In fitting the time- and azimuth-averaged pair production rate, we must fit the location of the boundary layer. Various techniques for defining and tracking the extent of the jet have been explored in the literature (e.g., Narayan et al. 2012; Yuan et al. 2015; Mościbrodzka et al. 2016). We find that fitting the location of the jet–disk boundary for each model is not justified by the improvement in fit to the pair production rate, and we simply fix $a = 1/2 \, \mathcal{L}$ and $b = 3 \, \mathcal{L}$ in Equation (6.28). Similarly we use $\sigma = 0.5$ and $\sigma_f = 0.1$ in Equation (6.27) for all models.

The parameters are likely to depend on magnetization, but we cannot evaluate this dependence because we consider only SANE models. The parameters may also depend on numerical resolution and the dimensionality of the model, which we also cannot evaluate with

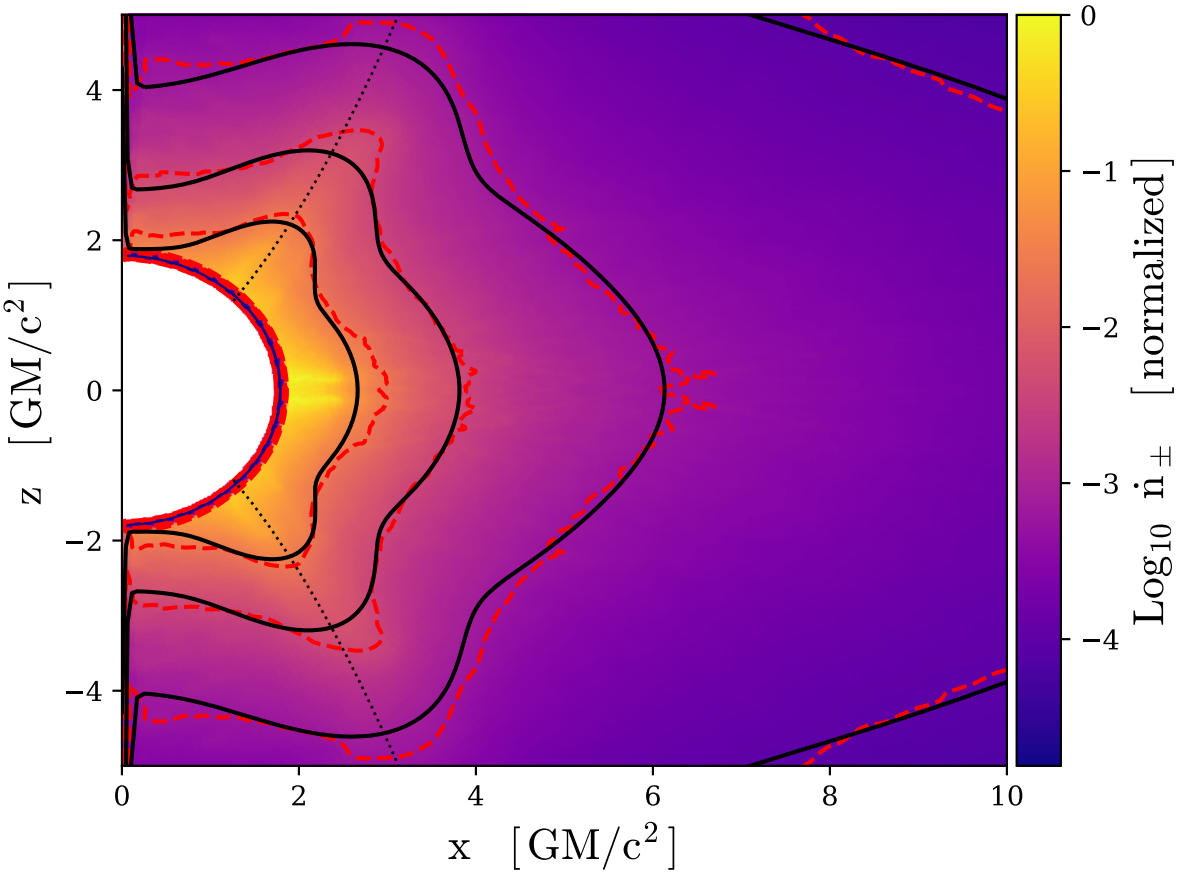


Figure 6.6: Pair production rate density (Model C). Numerically evaluated, time-averaged pair production rate density $\dot{n}_{\pm}$ as a function of position over domain for model C after vertical symmetrization over the disk midplane. Horizontal axis shows radial coordinate and vertical axis shows height above midplane. Solid colors correspond to $\log_{10}(\dot{n}_{\pm})$. Dashed red lines track contours in numerical value and solid black lines represent contours of model with fit parameters.

the existing model set. It is possible, for example, that the peak in the jet–disk boundary layer could increase as resolution increases and dissipation is concentrated in a narrower region within the boundary.

Figure 6.6 shows the time-and-azimuth averaged pair production rate for model B with fitting function contours overplotted. The fit is more accurate for inner regions of the disk where the $\dot{n}_{\pm}$ is large, but the fit works well even at larger radius.

In the low-$\dot{m}$ regime, we find

$$\mathcal{A}(m_8, \dot{m}) \approx 5.7 \times 10^{30}\ \dot{m}^{5.8}\ m_8^{-1.4} \tag{6.31a}$$

$$\mathcal{B}(m_8, \dot{m}) \approx 5.4 \times 10^{-4}\ \dot{m}^{-4/5} \tag{6.31b}$$

$$\mathcal{C}(m_8, \dot{m}) \approx 1.9 \times 10^{15}\ \dot{m}^{4}\ m_8^{-2} \tag{6.31c}$$

$$\alpha \approx 4.9\ \dot{m}^{-0.04} \tag{6.31d}$$

$$\kappa_1 \approx 2 \tag{6.31e}$$

$$\kappa_2 \approx 8. \tag{6.31f}$$

Again in the low-$\dot{m}$ regime, the total pair creation rate, integrated over the entire simulation domain, is well fit by

$$\dot{N}_{\pm}(m_8, \dot{m}) = 3.4 \times 10^{64}\ \dot{m}^{5}\ m_8^{1.5}. \tag{6.32}$$

Pairs are born with a broad spectrum of energies. The Lorentz factor of each lepton $\gamma_{\rm FF}$ as measured in the plasma fluid frame u_μ can be computed from the p^μ of the interacting photons since momentum is conserved. In the jet, the average pair is created with $\gamma_{\rm FF} \approx 10$. This result is consistent with that of M11.

6.6 Discussion

We have modeled drizzle pair production in simulations of SANE (low magnetic flux) black hole accretion flows in the mildly radiatively efficient regime and for select models corresponding to M87. The accretion simulations we consider model electron thermodynamics and radiative processes. Our models differ from M11 in several respects. First, electron heating is treated using the Howes (2010) model for dissipation at the bottom of a turbulent cascade. This model partitions dissipation approximately equally between electrons and ions when $B^2/(8\pi) \gtrsim P_{\rm gas}$ and preferentially heats the ions otherwise. Second, Coulomb coupling between ions and electrons is included. This transfers energy from ions to cooler electrons

and is a significant source of electron heating near the midplane at small radius (Ryan et al., 2017). Third, we self-consistently treat the transfer of momentum and energy between the plasma and the radiation field using the `ebhlight` code (Ryan et al., 2019). Finally, we consider bremsstrahlung emission when estimating the pair production rate (but not in the radGRMHD simulation, where it is energetically subdominant). Since bremsstrahlung produces a large population of photons with $h\nu \sim kT \sim 10 m_e c^2$, it can be important for pair production.

Our results largely agree with the analysis presented in M11. Still, there are interesting new questions we can answer. First: motivated by a new understanding of M87 based on the EHT 2017 results, is M87 likely to have a charge-starved magnetosphere? Second, are there differences in the geometry of pair production between state-of-the-art models and M11's more simplified treatment of electron thermodynamics?

6.6.1 Drizzle versus gaps

Pair drizzle can prevent the black hole magnetosphere from becoming charge starved and thereby forestall the opening of gaps and the generation of pair cascades. To see this, we compare the total number of available charges from both pairs and plasma $n_{\rm avail} \equiv n_\pm + n_{\rm pl}$ to the Goldreich–Julian charge density drawn from Equation (6.24), which uses the normal observer frame in its calculation. Using the characteristic time $\mathcal{T} \equiv \mathcal{L}/c$, we set $n_\pm \approx \dot{n}_\pm \mathcal{T}$. At $r = 2GM/c^2$ and along the pole at $\theta = 0$ (where $n_{\rm avail}$ is small since $n_{\rm pl}$ is negligible), the ratio is

$$\begin{aligned}\frac{n_{\rm avail}}{n_{\rm GJ}} &\approx 4.4 \times 10^{17}\, \dot{m}^{7/2}\, m_8^{3/2} \times \\ &\quad \left(1 + 1.1 \times 10^{12}\, \dot{m}^{6/5} + 9.9 \times 10^{14}\, \dot{m}^2\right). \end{aligned} \tag{6.33}$$

The terms in parentheses correspond to pair production by bremsstrahlung, Comptonized photons from the jet–disk boundary, and Comptonized photons from the midplane. Evi-

dently for $10^{-10} < \dot{m} < 10^{-4}$ the Comptonized boundary photons dominate; for $\dot{m} \gtrsim 10^{-7}$ the midplane photons dominate bremsstrahlung; and the midplane becomes increasingly important as $\dot{m}$ increases. The midplane may be more important than the boundary as $\dot{m} \gtrsim 10^{-3.7}$, but that extrapolates beyond the range of validity of our models.

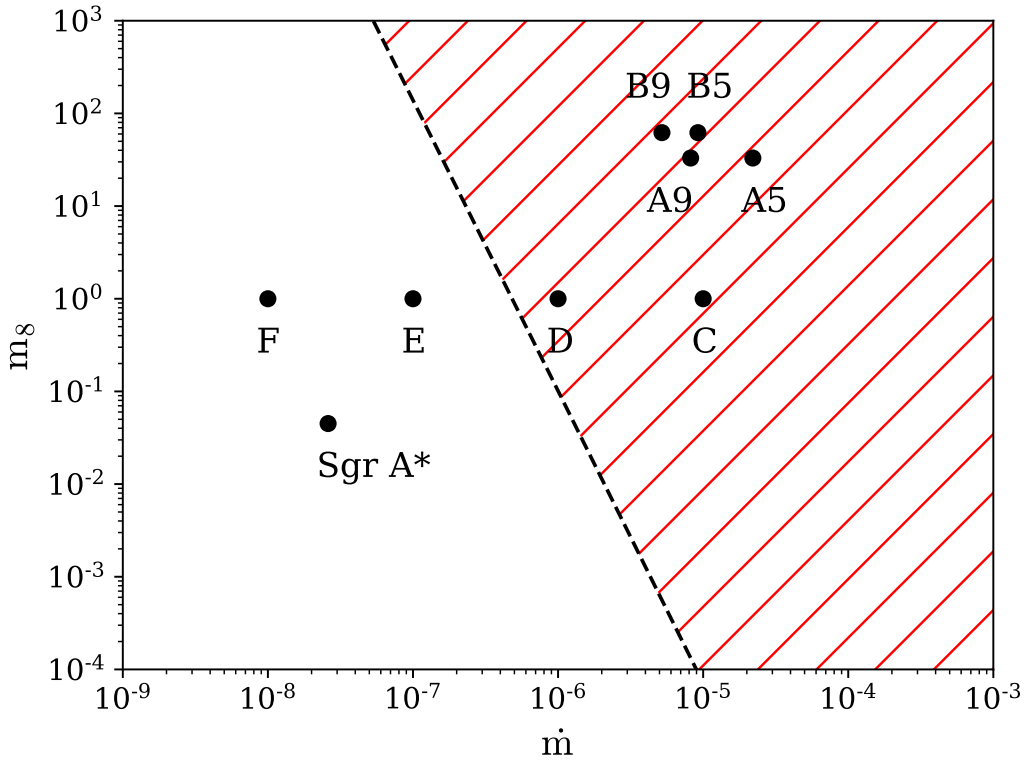


Figure 6.7: SANE $a_* = 0.5$ and $a_* = 0.94$ models in the $\dot{m}, m_8$ plane. The red hash marks show regions where the ratio of Goldreich–Julian density to the radGRMHD number density is below unity. In the unhatched region, the MHD approximation is not self-consistent.

In regions where $n_{\rm avail}/n_{\rm GJ} < 1$, the MHD approximation is not self-consistent. It seems likely that the outcome is a pair cascade (although this is not computable in our model) that increases $n_{\rm avail}$ by drawing on the free energy of the electromagnetic field and the radiation field until $n_{\rm avail}/n_{\rm GJ} \sim 1$. Figure 6.7 shows where in parameter space, according to Equation (6.33), the MHD approximation is not self-consistent. Figure 6.8 maps $n_{\rm avail}/n_{\rm GJ}$ in the poloidal plane for models C and E.

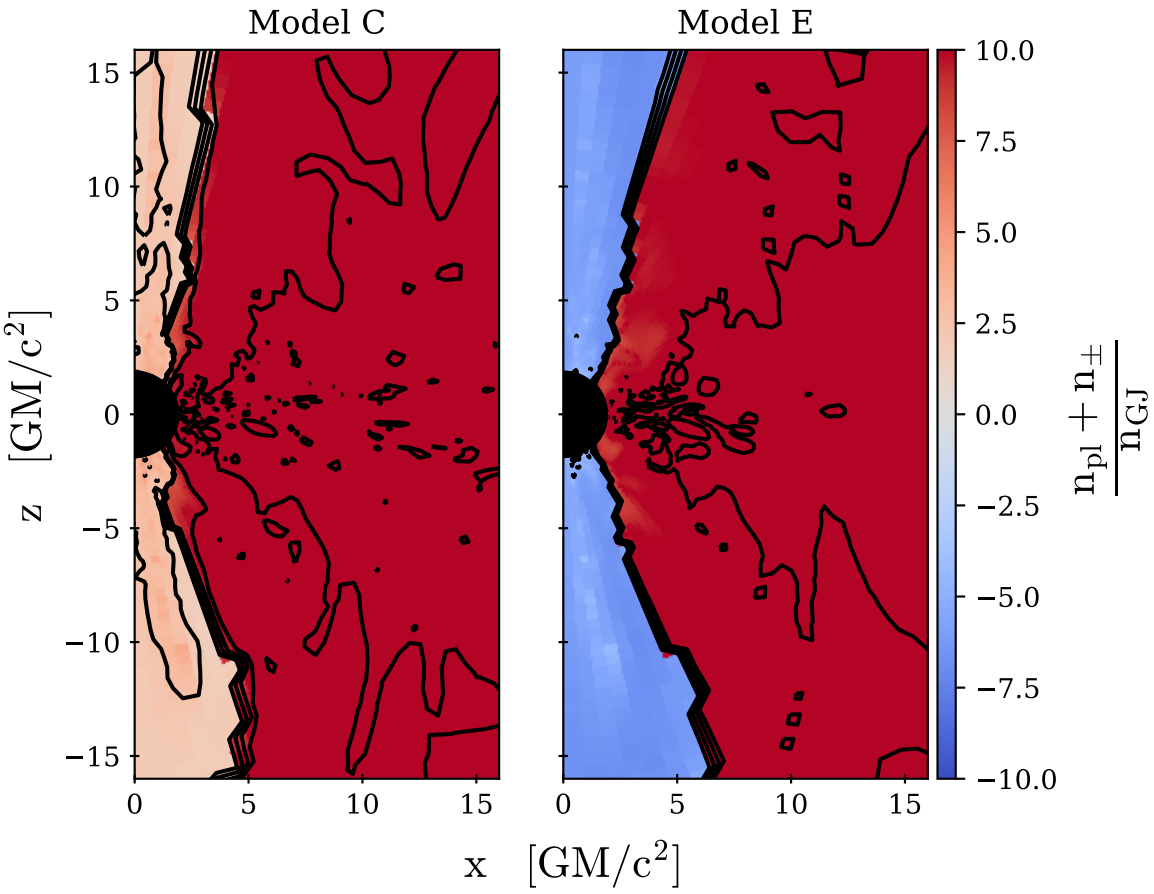


Figure 6.8: Ratio of available charge to Goldreich–Julian density (Equation (6.33)) for Models C and E ($a_* = 0.5, m_8 = 1$ with $\dot{m} = 1.1 \times 10^{-5}$ and 1.3×10^{-7} respectively). Black contours are evenly spaced in the log of the ratio. The black circle is the event horizon. Evidently the ratio is well above unity in the disk in both models, while the ratio in model E in the jet is far below unity and the MHD approximation is not self-consistent. Although pair cascades are not included in our model, they would appear difficult to initiate anywhere in model C, but they may be likely to occur in the jet region of model E.

6.6.2 Drizzle pair production power

As a black hole spins, it drags spacetime and the magnetic field lines near the horizon with it. These field lines produce an outward Poynting energy flux as they wind around the pole, via the BZ mechanism (Blandford & Znajek, 1977). The BZ mechanism is a favored explanation

for the source of black hole jet power. The BZ luminosity is given by

$$L_{\rm BZ} \equiv \int\limits_{\mu^2>\mu_f^2} T^r_{\ t\ {\rm EM}} \sqrt{-g}\ {\rm d}\theta\, {\rm d}\phi, \tag{6.34}$$

where $T^r_{\ t\ {\rm EM}} = b^2 u^r u_t - b^r b_t$ is the radial energy flux for the electromagnetic component of the stress-energy tensor. We find that the numerically computed values of $L_{\rm BZ}$ for our simulations match the fit given by Equation (36) of M11,

$$L_{\rm BZ} \approx 8 \times 10^{45} \left(1 - \sqrt{1 - a_*^2}\right)^2 \dot{m}\, m_8 \ \ {\rm erg\ s^{-1}}, \tag{6.35}$$

when $a_* = 0.5$.[7]

We now ask what fraction of the jet power can be accounted for by drizzle pairs. Using Equations (6.30) and (6.32),

$$\frac{L_\pm}{L_{\rm BZ}\Gamma_{\rm jet}} \approx 3.8 \times 10^{14}\ \dot{m}^4\, m_8^{1/2}. \tag{6.36}$$

Scaling Equation (6.35) to M87 using EHT results (Event Horizon Telescope Collaboration et al., 2019e) and assuming that $a_* \simeq 0.5$, $\dot{m} \simeq 10^{-5}$, and $m = 6.5 \times 10^9$, then $L_\pm \simeq 3 \times 10^{36}\ \Gamma_{\rm jet}$. In order for the drizzle pair luminosity to be comparable to M87's X-ray luminosity $\approx 10^{42}\,{\rm erg\,sec^{-1}}$, the typical pair would have to be born with an exceedingly high Lorentz factor $\Gamma_{\rm jet} > 10^6$. Thus, although drizzle-produced pairs may become important at $\dot{m} \sim 10^{-4}$ Eddington (suggested by Equation (6.36), but outside our model space), they account for a small fraction of total BZ power.

[7]Perturbative calculations of the BZ luminosity (to higher orders in the hole frequency) have been computed and compared to numerical simulation by, e.g., Tanabe & Nagataki (2008) and Tchekhovskoy et al. (2010). Tchekhovskoy et al. (2011) provided a generalized formula similar to the one given by M11 that also accounts for different magnetic fluxes near the horizon and thus treats both SANE *and* MAD accretion states.

6.6.3 Variability and the radiation model

Our fits for $\dot{n}_{\pm}$ and $\dot{N}_{\pm}$ represent the time- and azimuth-averaged behavior of the background $\gamma\gamma$ pair production process. In contrast, the instantaneous $\dot{n}_{\pm}$ does not peak along the entire boundary layer at once, but rather inside isolated island-like structures or plasmoids. The plasmoids are elongated in the radial direction, extend several M in width, and tend to travel along the boundary and evolve on timescales comparable with the dynamical time (see Nathanail et al. 2020 and Ripperda et al. 2020 for a discussion of plasmoids in nonradiative models). The plasmoid evolution depends strongly on the model parameters and the electron thermodynamics.

The highly variable plasmoid emission is dominated by Comptonized synchrotron photons rather than bremsstrahlung, which is generated mainly in the midplane at large radius and is relatively steady. Thus, in general, the high variability we observe in the domain-integrated pair production rate $\dot{N}_{\pm}$ is due to the rapid evolution of the plasmoids.

The scaling relations provided above have a limited range of validity. For $\dot{m} \gtrsim 10^{-5}$, radiative cooling is strong enough to qualitatively change the electron temperature distribution and thus the distribution of pair-producing photons. We are currently unable to explore this behavior because of the increasing computational intractability of running Monte Carlo radGRMHD simulations as optical depths to photon scattering increase and cooling times decrease relative to the light-crossing time of the domain.

6.6.4 Limitations and self-consistency

Because our model only considers pair production in a post-processing step, it cannot account for any back-reaction of drizzle pairs on the radiation field or underlying fluid dynamics. To check the self-consistency of this approximation, we can estimate the compactness parameter

$$l_c \equiv \frac{L_\gamma}{\mathcal{L}} \frac{\sigma_T}{m_e c^3}, \tag{6.37}$$

which is proportional to the optical depth to pair production. When $l_c \gg 1$, pair production cannot be treated as a perturbative process. In our models l_c ranges from 10^{-7} to 10^{-1} (see Table 6.1), so our treatment is self-consistent.

We assume that the electron distribution is purely thermal; however, since the plasma is collisionless everywhere in all our models (the Coulomb scattering mean free path is large compared to GM/c^2), the plasma need not fully relax to a thermal distribution. Moreover there is evidence for nonthermal electrons in both observations of low $\dot{m}$ accreting black holes (e.g., near-IR emission in the case of Sgr A*) and in simulations of collisionless, turbulent plasmas (e.g., Kunz et al., 2016). The presence of nonthermal electrons in a high-energy tail can result not only in higher-energy synchrotron photons, but also in an increase in Compton scattering events that increase photon energies to above the pair-producing threshold.

Our models used the Howes (2010) prescription for heating due to dissipation, in which the electron heating is driven by a Landau-damped turbulent cascade process. Other prescriptions (e.g., Rowan et al., 2017; Werner et al., 2018; Kawazura et al., 2019) would naturally produce a different electron temperature distribution. Because the pair luminosity depends strongly on electron temperature, modifications to the electron thermodynamics could significantly alter our results in ways that are difficult to assess without rerunning the radGRMHD models.

We computed pair production rates in post-processing using the fast-light approximation, in which it is assumed that the fluid does not change appreciably over the time it takes for light to travel across the simulation domain. It is possible but computationally expensive to dispense with this approximation (*slow light*). Performing a full slow light calculation would undoubtedly alter the pair production rate density on small length- and time-scales, but notice that the total radiative energy budget is conserved in both fast- and slow-light treatments, and so unless the fast-light approximation dramatically changes $\langle n_\gamma^2 \rangle / \langle n_\gamma \rangle^2$, the time-averaged pair production rate should not change significantly.

Finally, our models were limited to moderate resolution and two dimensions because of

the computational expense of running full radGRMHD simulations. Increasing resolution and especially performing simulations in three dimensions could change the profile of the jet–disk boundary layer and alter the dynamics of the plasmoid hotspots that develop within it. Since drizzle pair production peaks near the hotspots and is strongly dependent on the plasma temperature, the structure of $\dot{n}_{\pm}$ may change significantly with increased resolution or in the case of fully three-dimensional simulations.

6.7 Conclusion

We have modeled pair production due to the collision of photons in the background radiation field (here referred to as drizzle pair production) for sub-Eddington black hole accretion systems in the SANE state. Our plasma model is based on radGRMHD simulations using the `ebhlight` code (Ryan et al., 2017, 2018), which evolves the plasma and the full energy-dependent photon distribution. The radGRMHD evolution includes synchrotron emission, absorption, and Compton scattering. It also separately evolves ion and electron internal energies and explicitly accounts for dissipation using the electron heating prescription of Howes (2010). We post-processed the fluid data using Monte Carlo radiation transport to track pair production due to photon–photon collisions. In the post-processing we included bremsstrahlung emission, which is a potentially important source of photons near the pair production threshold.

Our approach closely follows M11 and extends it in several ways. We use energetically self-consistent radGRMHD models rather than nonradiative GRMHD models. We incorporate a dissipation model rather than fixing the ion-to-electron temperature ratio. We study multiple black hole spins ($a_* = 0.5$ and 0.94). Finally, we include bremsstrahlung emission, which is non-negligible at frequencies near the pair-production threshold.

Our key findings are:

1. The importance of cooling increases as accretion rate increases. This leads to a

shallower dependence of the source-integrated pair production rate on $\dot{m}$ than in M11.

2. The spatial distribution of pair production peaks within the jet–disk boundary, in contrast to M11. This is because electron temperature peaks in the boundary layer, and drizzle pair production closely follows the electron temperature profile. These results are summarized by Equations (6.27), (6.29), and (6.32), with the parameter values reported in Equation (6.31).

3. The pair production rate density can be divided into spatially distinct bremsstrahlung and Comptonized synchrotron components. The bremsstrahlung component is comparatively steady and lies in the midplane, outside the midplane Comptonized component. The bremsstrahlung component is weaker than the Comptonized component for all models considered here.

4. The Comptonized component from the midplane becomes comparatively larger as $\dot{m}$ increases, but it is dominated by the boundary layer Comptonized component for all models considered in this paper.

5. The drizzle pair production rate is time variable, with the difference between subsequent samples occasionally approaching four orders of magnitude. These variations are dominated by fluctuations in the synchrotron and Compton components of the background radiation field within the jet–disk boundary.

6. We confirm the finding of M11 that the drizzle process (in M87-like SANE models) produces a background pair density that is far above the Goldreich–Julian density. This suggests that it will be difficult to open gaps absent some dynamical process that is not incorporated in our models.

7. We confirm the finding of M11 that drizzle pair production in Sgr A*-like SANE models is too feeble to keep the pair density above the Goldreich–Julian density. In GRMHD models

the expected BZ power is

$$\sim 2.8\, a_*^2(\phi/15)^2 \dot{M}c^2 = 1.6 \times 10^{38}\, a_*^2(\phi/15)^2(\dot{M}/(10^{-9} M_\odot \mathrm{yr}^{-1})\, \mathrm{erg\,sec}^{-1} \tag{6.38}$$

. The difficulty in firmly identifying a jet with comparable power suggests it is not present. Parfrey et al. (2019) PIC-based magnetosphere model suggests that this is *not* due to a fundamental change in the BZ power for charge-starved magnetospheres.

In future work we plan to explore drizzle pair production in MAD models where the increased electron temperatures and magnetic field strengths provide a more favorable environment for high-energy photons production. We also plan to extend the calculations to three dimensions, which will provide the opportunity to study transient behavior associated with the characteristic nonaxisymmetry of MAD models.

Chapter 7

Black hole glimmer

Gravitational lensing near a black hole is strong enough that light rays can circle the event horizon multiple times. Photons emitted in multiple directions at a single event, perhaps because of localized, impulsive heating of accreting plasma, may take multiple paths to a distant observer. In the Kerr geometry, each path is associated with a distinct light travel time and a distinct arrival location in the image plane, producing what I call *black hole glimmer*. This glimmer sequence of arrival times and locations uniquely encodes the mass and spin of the black hole and can be understood completely in terms of properties of bound photon orbits, separating it from the details of the accretion model and the messy gastrophysics. In this chapter, I provide a geometrically motivated treatment of Kerr glimmer and evaluate it numerically for simple hotspot models to show that glimmer can be measured in a finite-resolution observation.

7.1 Introduction

Spinning supermassive black holes likely power relativistic jets via the Blandford–Znajek mechanism (Blandford & Znajek 1977, although see Blandford & Payne 1982 and Lynden-Bell 2006 for alternative jet-power mechanisms). To probe the spin–jet connection through observation, it is necessary to understand the properties of spacetime near the hole and to have an accurate model of the accreting plasma around the hole. If spacetime is described by the Kerr metric, then its properties are uniquely determined by the mass and angular

This chapter is lightly adapted from G. N. Wong, ApJ, Volume 909, Issue 2, id.217.

momentum of the central black hole.

Many spin measurement methods propose measuring the size of the accretion disk, the qualities of the disk oscillations, or the deviation in line profiles due to black hole spin (e.g., Hanawa, 1989; Kojima, 1991; Laor, 1991; Kato, 2001; Miller, 2007). These measurement techniques rely on an accurate understanding of the accretion flow and are therefore subject to uncertainties in the plasma physics model. The 2017 Event Horizon Telescope observation of the black hole at the center of the galaxy M87 provided the first direct horizon-scale observation of a black hole (Event Horizon Telescope Collaboration et al., 2019a); however, its ability to constrain the spin of the hole is also limited by the modeling uncertainties.

Strong gravitational lensing allows photons to travel along bound orbits that circle black holes (see Claudel et al. 2001 for a treatment of bound photon surfaces in arbitrary spacetimes). In the Kerr metric, the properties of these orbits are determined solely by the spacetime geometry, i.e., by the mass and angular momentum of the black hole, in the case of the vacuum metric. In the image plane, the asymptotically bound orbits produce a characteristic *critical curve*, whose size and shape are also set by the spacetime geometry. Since the critical curve is determined solely by the spacetime, it is independent of the accretion model; thus it provides a consistent signature that directly probes the hole's properties.

Although the critical curve is not necessarily observable, it is often traced by a high-intensity *photon ring* that is produced as nearly bound orbits steadily sample the high-emissivity region near the hole (see Gralla & Lupsasca 2020 for a discussion of some differences between the signatures of the curve and ring, especially in the case of midplane emission). Measurement strategies to infer bounds on spin from the shape of the ring or curve have been proposed in the past (e.g., Falcke et al., 2000; Takahashi, 2004; Bambi & Freese, 2009; Hioki & Maeda, 2009; Younsi et al., 2016; Johnson et al., 2020). Related proposals have suggested testing the hypothesis that spacetime is Kerr by looking for deviations in the shape of the curve (e.g., Amarilla et al., 2010; Tsukamoto et al., 2014; Amarilla & Eiroa, 2013; Mizuno et al., 2018; Medeiros et al., 2020; Olivares et al., 2020; Wielgus et al.,

2020).

The bound orbits allow light near a black hole to orbit it multiple times before escaping to an observer. Since emitters can radiate in multiple directions simultaneously, two rays produced by the same source may orbit the hole a different number of times. Light signals from subsequent orbitings will be separated by a time delay set by the length of a complete winding around the hole. These delays will cause the source to echo in the image plane. Since the echo period is a function of path length, it is intrinsically tied to the underlying bound orbits and can be measured to infer spin. Constraining spin by measuring dominant echoes has been considered in the past (e.g., Broderick & Loeb, 2005; Moriyama & Mineshige, 2015; Saida, 2017; Thompson, 2019; Moriyama et al., 2019; Gralla & Lupsasca, 2020).

In detail, the Kerr geometry produces a rich spectrum of echo time delays associated with resonant bound orbits. These resonant echo delays are closely related to the black hole quasinormal-mode spectrum in the eikonal limit (see Yang et al., 2012), which has been studied in the context of gravitational-wave ringdown and measuring mass and spin (e.g., Berti et al., 2006; Buonanno et al., 2007; Berti et al., 2007).

Each echo maps to a distinct arrival location in the image plane; taken together, the set of echoes produces a characteristic *black hole glimmer* that encodes the mass and spin of the hole. Since the mechanism that produces these echoes is driven purely by the spacetime, the black hole glimmer signature is separable from the source emission model. If glimmer can be measured precisely, it is possible to test the Kerr hypothesis and infer the black hole mass and spin, even without a detailed understanding of the emission source. We provide a geometrically motivated treatment of Kerr glimmer and demonstrate that the glimmer signature of a hotspot can be measured even in a finite-resolution observation.

This paper is structured as follows. In Section 7.2, we review the salient features of the Kerr geometry in the context of bound orbits, and in Section 7.3 we describe the static observable image due to the bound orbits. We describe and explore the Kerr glimmer signature in Section 7.4, and we give a brief discussion and example measurement in Section 7.5. We

summarize in Section 7.6.

7.2 Kerr geometry and bound orbits

Black holes (in vacuum) in general relativity are described by their mass, angular momentum, and charge, although it is unlikely that supermassive black holes with a dynamically important charge exist in nature. Charge-neutral holes with nonzero angular momentum are described by the Kerr metric. We use geometrized units with $G = c = 1$ and write angular momentum J in terms of the conventional dimensionless spin parameter $a_* \equiv J/M^2$ where M is the mass of the hole. Hereafter, we set $M = 1$. In Boyer–Lindquist coordinates $x^\mu = (t, r, \theta, \phi)$, the Kerr line element is (Bardeen et al., 1972)

$$\begin{aligned} \mathrm{d}s^2 = &- \left(1 - \frac{2r}{\Sigma}\right) \mathrm{d}t^2 - \frac{4\, a_* r \sin^2\theta}{\Sigma}\, \mathrm{d}t\, \mathrm{d}\phi + \frac{\Sigma}{\Delta}\, \mathrm{d}r^2 \\ &+ \Sigma\, \mathrm{d}\theta^2 + \frac{\left(r^2 + a_*^2\right)^2 - \Delta a_*^2 \sin^2\theta}{\Sigma} \sin^2\theta\, \mathrm{d}\phi^2 \end{aligned} \tag{7.1}$$

with

$$\Sigma \equiv r^2 + a_*^2 \cos^2\theta, \qquad \Delta \equiv r^2 - 2r + a_*^2\,. \tag{7.2}$$

The Kerr geometry admits a set of spherical bound orbits—orbits with fixed radial coordinate—for both massive and massless particles near black holes. The set of all spherical massless/photon orbits is known as the *photon shell* (Johnson et al., 2020). The observational consequences of the photon shell on black hole images have been studied both in the context of nonspinning holes (see especially Luminet, 1979) and the more general spinning Kerr hole (e.g., Bardeen, 1973). One of the first full treatments of spherical photon orbits in Kerr was presented by Teo (2003), who also provided a convenient categorization of the types of such orbits. In this section, we review several key features of the bound photon orbits. § 7.7 details the results described in this section. We do not consider black holes

with extremal spin.

For a given spin, the set of all bound spherical photon orbits can be parameterized by the radii of the orbits. These radii lie continuously in the range $r_- \leq r \leq r_+$, where

$$r_\pm = 2\left(1 + \cos\left(\frac{2}{3}\cos^{-1}\pm a_*\right)\right). \tag{7.3}$$

Only the two extremal orbits at $r_\pm$ are confined to the midplane. The prograde orbit at $r = r_-$ revolves around the hole in the same direction as its spin, and the retrograde orbit revolves opposite the spin at $r = r_+$. The other orbits at intermediate radii oscillate between symmetric minimum and maximum latitudes $\theta_\pm$.

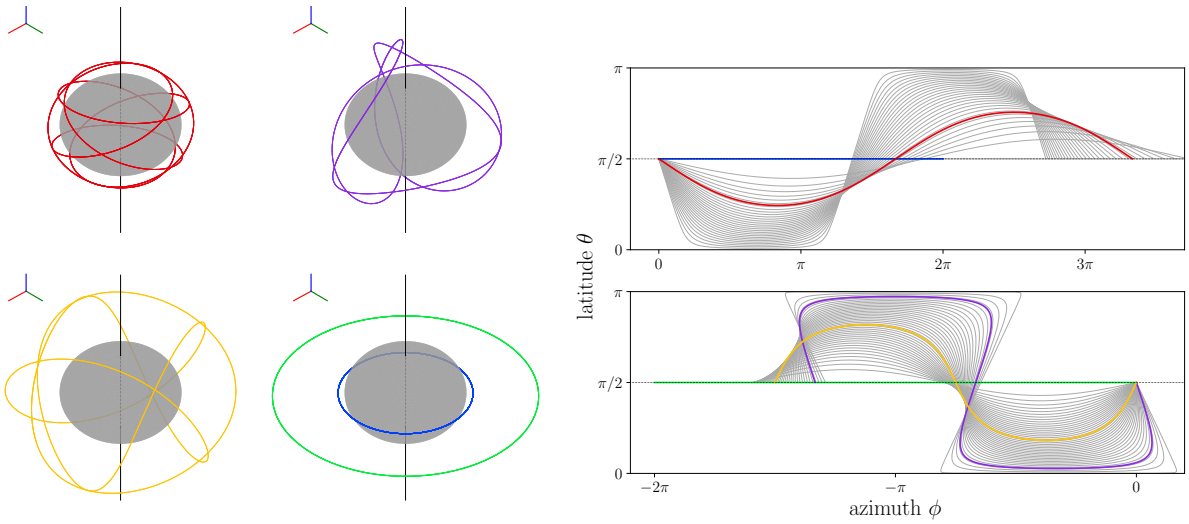
For holes with nonzero spin, the latitudinal oscillation period is different from the azimuthal ϕ period, so after a full ϕ orbit around the hole ($\phi \to \phi + 2\pi$), the geodesic will not return to the same θ coordinate. The magnitude of the precession can be written in terms of the deviation in ϕ from one latitudinal cycle to the next. In the case of no precession, $\Delta\phi$ would be 2π. In general, it is given by

$$\begin{aligned}\Delta\phi &= 4\int_0^{\theta_+} \frac{\mathrm{d}\phi}{\mathrm{d}\theta}\,\mathrm{d}\theta \\ &= \frac{4}{\sqrt{-u_-^2}}\left(\frac{2r - a_*\Phi}{\Delta}K\left(\frac{u_+^2}{u_-^2}\right) + \frac{\Phi}{a_*}\Pi\left(u_+^2, \frac{u_+^2}{u_-^2}\right)\right),\end{aligned} \tag{7.4}$$

where K and Π are complete elliptic integrals[1] of the first and third kinds, the roots of the latitudinal potential are

$$u_\pm^2 \equiv \frac{a_*^2 - Q - \Phi^2 \pm \sqrt{\left(Q + \Phi^2 - a_*^2\right)^2 + 4a_*^2 Q}}{2a_*^2}, \tag{7.5}$$

[1]We use the square of the elliptic modulus as the parameter for all elliptic integrals in this work.



(a) Resonant orbits in 3d

(b) Bound orbit trajectories

Figure 7.1: Left panel: five selected resonant orbits for a spin $a_* = 4/5$ hole. Right panel: trajectory of bound orbits for the same $a_* = 4/5$ hole plotted in the θ–ϕ plane. The resonant orbits in the left panel are colored in the right panel. Here, $\Delta\phi$ for each orbit corresponds to the ϕ displacement after one complete latitudinal cycle, i.e., when the trajectory ends on the plot. The spread in $\Delta\phi$ across the different orbits increases with the spin a_*.

and the constants of motion $\Phi(r)$ and $Q(r)$ are given by

$$\Phi = -\frac{r^3 - 3r^2 + a_*^2 r + a_*^2}{a_* (r-1)} \tag{7.6}$$

$$Q = -\frac{r^3 (r^3 - 6r^2 + 9r - 4a_*^2)}{a_*^2 (r-1)^2}. \tag{7.7}$$

Figure 7.1 plots the trajectory of bound orbits in the θ–ϕ plane and illustrates this precession effect, where one complete latitudinal cycle does not correspond to an azimuthal displacement of 2π. The spread in $\Delta\phi$ increases with a_*.

By studying $\Delta\phi$, we can identify which orbits are also closed—which geodesics return to their original positions and orientations. Closed (resonant) orbits correspond to the case that $\Delta\phi/2\pi$ is a rational number $= p/q$ in simplest form. Infinitely many closed orbits exist, but orbits with small p, q are the most interesting here, since they have the shortest path lengths. Figure 7.1 shows the trajectories of several closed orbits both in three dimensions

and as projected on the latitude–azimuth plane.

We also compute the time delay for one complete latitudinal cycle of the orbit,

$$\begin{aligned}\Delta t &= 4\int_0^{\theta_+} \frac{\mathrm{d}t}{\mathrm{d}\theta}\,\mathrm{d}\theta \\ &= \frac{\left(r^2+a_*^2\right)^2 - 2a_*\Phi r - a_*^2\Delta}{a_*\Delta\sqrt{-u_-^2}} K\left(\frac{u_+^2}{u_-^2}\right) \\ &\quad - 4a_*\sqrt{-u_-^2}\left[K\left(\frac{u_+^2}{u_-^2}\right) - E\left(\frac{u_+^2}{u_-^2}\right)\right],\end{aligned} \tag{7.8}$$

where E is the complete elliptic integral of the second kind. In the case of closed orbits, the time delay to complete a full cycle is a function of the number of latitudinal oscillations per complete cycle:

$$\Delta T = q\ \Delta t\ 2\pi/\Delta\phi. \tag{7.9}$$

7.3 The image of the photon shell

We now consider the signature of the photon shell as seen by an observer far away from the hole. Since the orbits comprising the shell are unstable, photons whose paths deviate slightly from the precise trajectories described above will either fall onto the hole or escape to infinity, where they can be captured by an observer. The exponential instability can be described in terms of how the deviation δr between a geodesic's radial position and the radial position of the corresponding bound orbit increases (or decreases) after n_ϕ azimuthal cycles around the hole,

$$\delta r(n_\phi) = \exp\left(\pm\gamma_\phi n_\phi\right)\delta r(0). \tag{7.10}$$

This equation defines the Lyapunov exponent γ_ϕ, which is given by

$$\gamma_\phi = \frac{2\pi}{\Delta\phi}\, 2\gamma_\theta \tag{7.11}$$

where γ_θ is taken to be consistent with Johnson et al. (2020) and governs the deviation after one latitudinal θ half-cycle (from midplane to extremum back to midplane),

$$\gamma_\theta = \frac{4}{a_* \sqrt{-u_-^2}} \sqrt{r^2 - \frac{r\Delta}{(r-1)^2}} K\left(\frac{u_+^2}{u_-^2}\right) \ . \tag{7.12}$$

This Lyapunov exponent, the time delay of Equation 7.4, and the azimuthal period of Equation 7.8 characterize strong lensing by Kerr black holes (Gralla & Lupsasca, 2020).

Rather than consider the source-to-observer model, in which we start with all geodesics that are emitted from a source and select only those that make it to the observer, it is convenient to switch to the observer-to-source model, where we begin with the set of all geodesics that intersect the image plane. For a camera at infinite distance, all photons incident on the camera arrive parallel to each other.

Following Bardeen (1973), we parameterize geodesics that intersect the image plane according to their impact parameters x and y (α and β in Bardeen). By convention, the y-axis is aligned with the projection of the black hole spin axis on the image. It is sometimes more convenient to work with polar coordinates on the image $\rho = \sqrt{x^2 + y^2}$ and $\varphi = \arctan y/x$ (see the top left panel of Figure 7.2).

Geodesics at large impact parameter (far from the image center) remain far from the hole and barely feel its influence. As the impact parameter decreases, however, the geodesics become increasingly bent, and eventually they wrap around the hole and undergo latitudinal oscillations. The set of impact parameters for all geodesics that undergo a fixed number of oscillations defines a subring on the image.

Nested subrings produce demagnified images of space. The magnification is given by the image area of the subring and scales according to the Lyapunov treatment above. If most

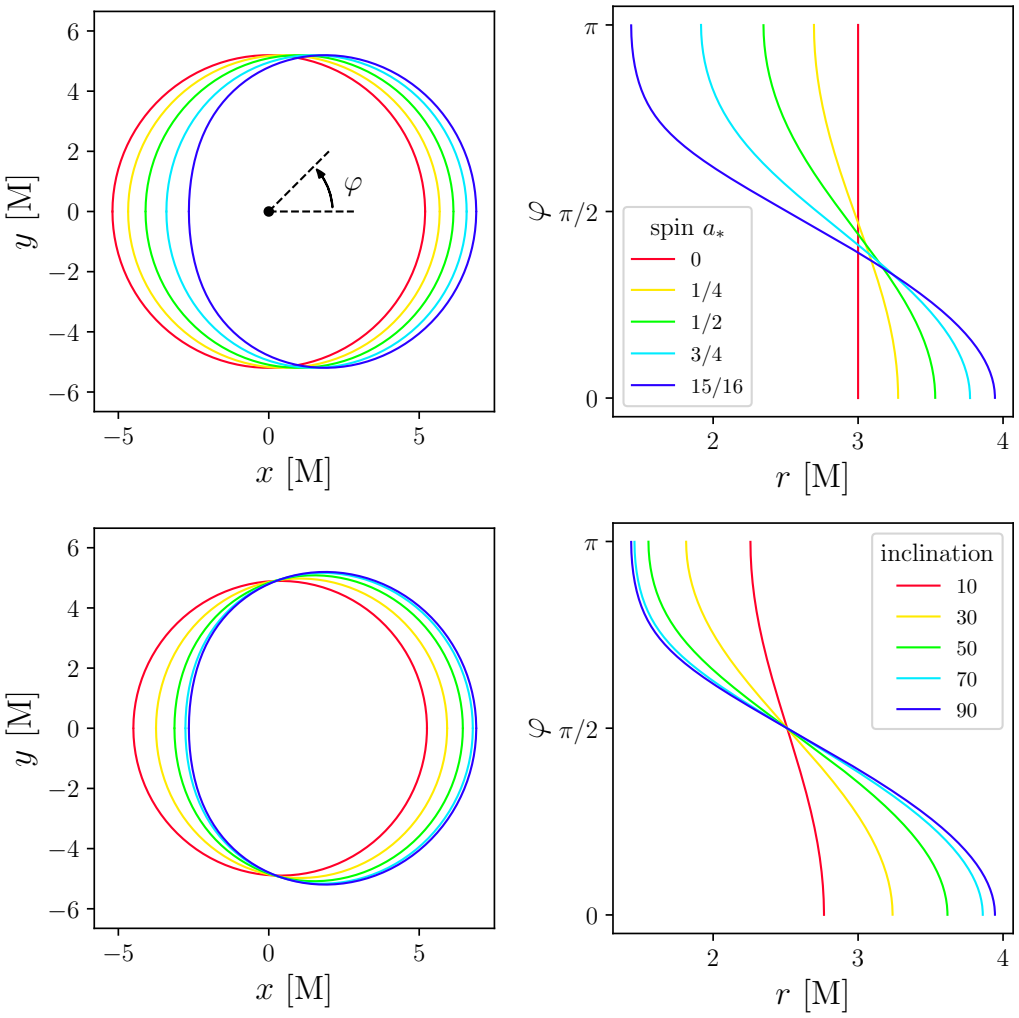


Figure 7.2: Left column: size and shape of critical curve on image for (top) black holes with different spin viewed edge-on or (bottom) black holes with $a_* = 15/16$ viewed at different inclinations. The top-level panel shows how the angle φ is measured counterclockwise from the positive x-axis. Right column: mapping between points along the critical curve φ and the Boyer-Lindquist radius of the probed bound orbit.

of the emission is produced near the black hole, then it will be sampled n times in the nth subring, so the ratio of flux (area-integrated intensity over the subring on the image) between the n and $n+1$ subrings will be given by

$$\frac{F_n}{F_{n+1}} \approx \frac{e^{-\gamma(n-1)} - e^{-\gamma n}}{e^{-\gamma n} - e^{-\gamma(n+1)}} = e^{\gamma}. \tag{7.13}$$

In the limit that n goes to infinity, geodesics wrap around the hole infinitely many times and are effectively trapped by the bound orbits described above. The set of all impact parameters with $n \to \infty$ defines the *critical curve* on the image. The region within the critical curve is sometimes called the black hole shadow.

Points on the critical curve correspond to different bound orbits—different radii—within the photon shell, so the path of the curve can be parameterized by r (see Bardeen 1973 and also Johannsen 2013):

$$x = -\frac{\Phi}{\sin i} \tag{7.14}$$

$$y = \pm\sqrt{Q + a_*^2 \cos^2 i - \Phi^2 \cot^2 i}. \tag{7.15}$$

The shape of the critical curve is thus a function of two parameters: the spin of the hole and the viewing inclination angle i. Figure 7.2 shows the shape of the critical curve both for black holes with different spins (top) and for the same hole observed at different inclinations (bottom). It also shows the mapping between points along the critical curve (parameterized by φ) and the bound orbits they correspond to (parameterized by r). The orientation of the curve is determined by the projection of the black hole spin axis on the image plane.

The spin of the hole determines which radii support bound orbits. Varying the inclination angle limits the set of bound orbits that are visible to the observer (for example, purely equatorial orbits are not accessible to a top-down observer). The set of accessible radii decreases from the full range given by Equation 7.3 at $i = 90°$ to the the single radius corresponding to $\Phi = 0$ at $i = 0°$ (for more detail, see § 7.7 and especially Equation 7.35).

7.4 The Kerr glimmer signature

Although messy gas dynamics determine the larger image features, the photon shell produces a separable, unique signature in the image domain. This signature is independent of the gas dynamics model, so it provides a direct way to measure the properties of the underlying black

hole. The presence of an infinite number of subrings on the image means that an (optically thin) emission source will be imaged an infinite number of times, albeit with exponentially decreasing flux from one instance to the next. The image in the nth subring comes from light that has gone around the hole n times, and each subring image will echo with a period equal to the total light travel time around the hole.

The aggregate signature produced by an emission source depends on the characteristics of the source, but we can provide an initial analysis by making two remarks in the context of a simplified model. First, since different positions along the critical curve correspond to different bound orbits and path lengths, the delay between subsequent imagings will be a function of position on the curve.

Second, if the source emission is localized in space, then an orbit that probes the source must return to the same localized area in order for an echo to be excited along that geodesic. Since bound orbits precess, a geodesic that passes through a fiducial source may take many revolutions around the hole before it passes through the source again. Since flux decreases exponentially with n_ϕ, the precession can render some orbits practically echoless. The latter criterion is the most restrictive and is relaxed in the second model we present.

In practice, the echo response function to an individual emission event is complicated, since it depends on the size and duration of the source. The details are complicated further by the initial transient response, which is determined by source size, duration, and position. Rather than attempt a full treatment of the emission variability near the hole, we consider simplified hotspot models. In Section 7.5, we argue that the hotspot model is not restrictive, especially in the limit of an optically thin source.

Each hotspot is taken to be a transient, isolated Gaussian blob with a time- and position-dependent emissivity,

$$j(t,\vec{r}) \propto \Psi(t) \exp\left(-\delta r^2/2\sigma_r^2\right) \exp\left(-\delta t^2/2\sigma_t^2\right), \tag{7.16}$$

where δr is the distance to a point on a Keplerian orbit at radius r_0, $t_0 \equiv t - \delta t$ sets the time when the hotspot is brightest, σ_r and σ_t describe the width of the hotspot in space and time, and where $\Psi(t)$ is a bump function in time

$$\Psi(t) = \begin{cases} \exp\left(\frac{1}{\delta t^2 - \sigma_t^2}\right) & |\delta t| < \sigma_t \\ 0 & \text{otherwise,} \end{cases} \tag{7.17}$$

which forces a smooth decay to zero emissivity.

7.4.1 Point-source emission

We start by considering the response produced by a unidirectional point-source emitter as measured along the critical curve. Since echoes occur when geodesics sample the same emission source multiple times, then in the limit as $\sigma_r \to 0$, the only universal echoes that occur at a fixed φ must come from closed bound orbits, since closed orbits are the only ones that return to the exact same point while following the exact same heading. Other (nonclosed) geodesics will either miss the emitter, not arrive at the correct observer location, or pass through the emitter with a different heading and thus accrue no new intensity. The potential for an arbitrary emission source to be imaged but not produce a trivial echo is what differentiates Kerr from Schwarzschild, since all spherical orbits in Schwarzschild are closed.

Even though an infinite number of closed orbits exist, only orbits with short path lengths will produce observable echoes since the sensitivity required to detect echoes from multiple turnings increases exponentially with the number of turnings according to the Lyapunov treatment summarized by Equation 7.13. Thus, for a given spin, there is a limited number of fixed, position-dependent observable echo time delays. Notice that although the signal from echoes with longer delays is more attenuated, the overall magnitude of the dimming may decrease as spin is increased.

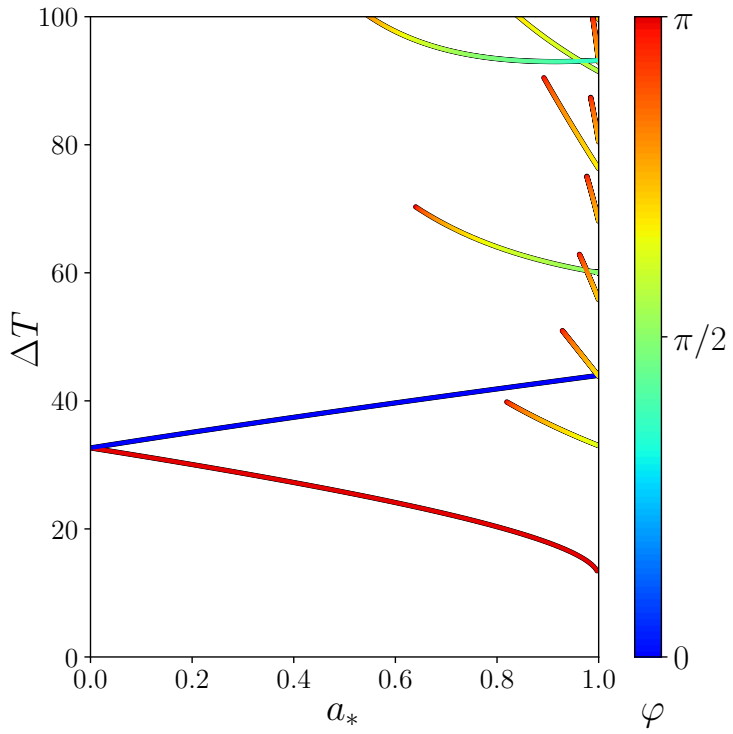


Figure 7.3: Echo delay times due to resonant orbits as a function of black hole spin. As spin increases, the number of accessible short-time-delay echoes increases. Color encodes the location of the echo on the critical curve. Only perfect resonances are shown in this plot.

Figure 7.3 plots the allowed echo delays due to closed orbits as a function of spin. Because each delay is due to a particular bound orbit, it maps to a distinct position along the critical curve. The allowed delays in the figure are colored according to the angle φ along the critical curve where they appear. The time delays for the prograde and retrograde orbits are always accessible, since they lie within the equatorial plane and thus always pass through the same points. The closed orbits can be computed by identifying which radii in Equation 7.4 correspond to (the reciprocals of) the first few levels of a Stern–Brocot tree. The nontrivial radial structure of the Kerr spacetime means that the time delay for a geodesic that circles the hole twice before closing will not be twice the delay for a geodesic that only circles it once, so the higher-order resonances are not perfect multiples of the lower-order ones.

Thus far, we have considered unidirectional emitters. If this restriction is removed, new echoes with different periods may be produced when a geodesic intersects itself along a dif-

ferent heading. For example, closed orbits with an even-to-odd ratio between the number of azimuthal and latitudinal cycles will return to the same point (with mirrored latitudinal heading) after having completed only half of the full closed path. Echoes due to self-intersections off of the midplane (see Figure 7.1 for examples) will lead to more complicated, composite delay structures that repeat with the same fundamental periodicity of the closed orbits. We return to this point in our discussion of hotspot position.

7.4.2 Finite-width sources

Real-world emission sources have nonzero width. In the context of glimmer, increasing the width of an emission sources softens the condition that a geodesic must return to the same point in (θ, ϕ) space for an echo to be produced, since geodesics will resample the fiducial source feature as long as the deviations in their positions are smaller than the size of the feature. These more permissive conditions broaden the set of accessible time delays and allow a wider range of φ to clearly echo.

The echo response can be visualized by plotting the intensity of light observed at different angular positions along the critical curve as a function of time. Black holes with different spins and observers at different inclinations will see echoes produced by bound orbits at different radii. Figure 7.4 shows how changing the spin or inclination affects the echo response produced by a Keplerian hotspot with $r_0 = 2.8\,\mathrm{M}$, $\sigma_r = 0.5\,\mathrm{M}$, and $\sigma_t = 1\,\mathrm{M}$. For each echo response, the radii that admit bound orbits (determined by spin) are identified by a black line, and the subset of radii with visible bound orbits (determined by inclination) is denoted by a red bar.

The top row of Figure 7.4 shows how the echo response changes as the angular momentum of the central black hole increases. As seen in Figure 7.3, the difference between the echo delays for the prograde ($\varphi = \pi$) and retrograde ($\varphi = 0, 2\pi$) orbits increases as spin increases. The bottom row of Figure 7.4 shows the effect of varying inclination for a black hole with fixed spin $a_* = 15/16$. At the top-down $i = 0°$ inclination, every point along the critical

curve corresponds to an orbit at the same $\Phi = 0$ radius (see Equation 7.35), so the echo delay period is constant around the curve. Notice that for $a_* = 15/16$, the $\Phi = 0$ orbit is not closed, so the echo blips drift around the curve. As inclination increases to $i = 90°$, the full set of bound orbits becomes visible.

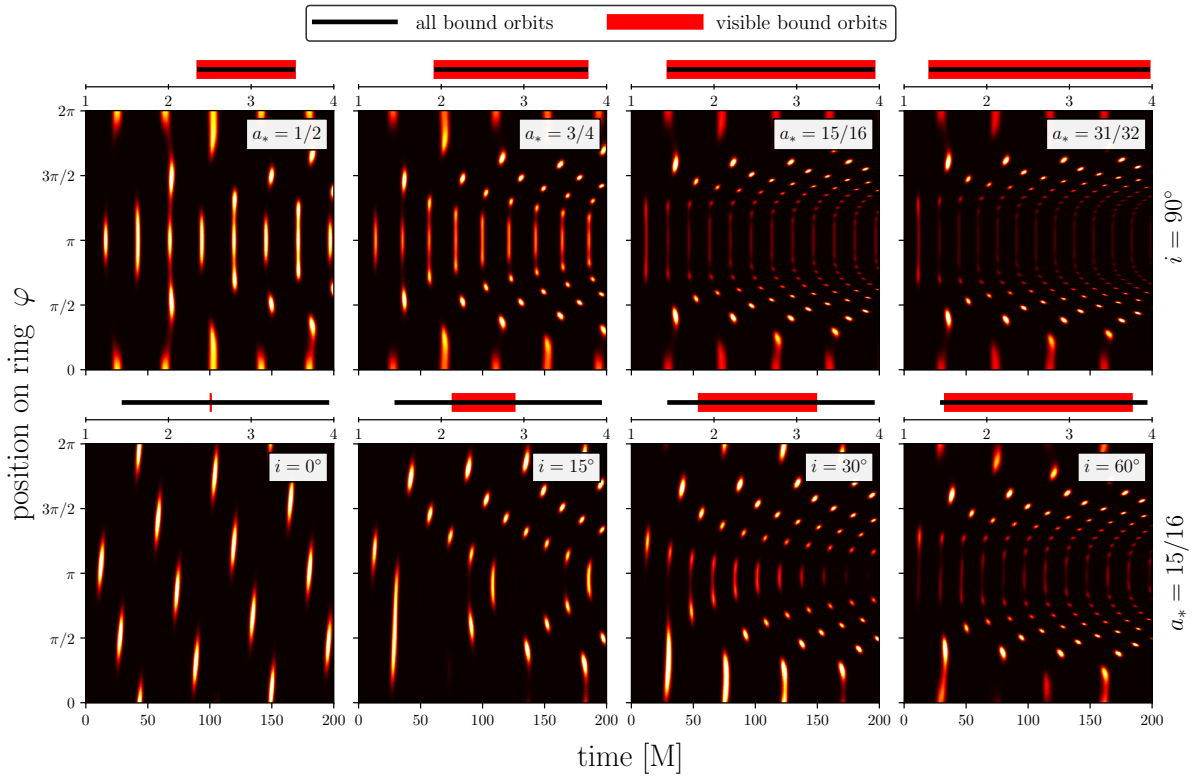


Figure 7.4: Echo response produced by a short-lived, Keplerian hotspot ($r_0 = 2.8\,\mathrm{M}$, $\sigma_r = 0.5\,\mathrm{M}$, $\sigma_t = 1\,\mathrm{M}$) for black holes with different spins and for observers at different inclinations. Color encodes the time-dependent intensity of light as a function of angle along the critical curve $\varphi \in (0, 2\pi)$. The bar above each panel shows the range of radii corresponding to bound orbits (black) and the set of visible bound orbits (red). Top row: the variation in the response as a function of black hole spin. The spread in echo delays between the prograde orbit ($\varphi = \pi$) and the retrograde orbit ($\varphi = 0, 2\pi$) increases as spin is increased. Bottom row: the echo response for the same $a_* = 15/16$ hole but observed at different inclinations. As inclination increases, the set of visible bound orbits increases, so more echoes become visible. Since the $\Phi = 0$ orbit is not closed for $a_* = 15/16$, the $i = 0°$ echo blips drift through φ. An animation showing the evolution of the echo response as a function of spin and inclination is available online.

The glimmer signature comprises a mapping between echo delay periods and positions

along the critical curve. Since this mapping is due solely to the spacetime geometry, it is independent of emission details like hotspot position, shape, and size. In contrast, the details of the full echo response, such as the relative strengths or phase offsets of echoes at different φ, are influenced by the characteristics of the hotspot.

Figure 7.5 shows the echo responses produced by different hotspots. In addition to plotting the echo response (as in Figure 7.4), Figure 7.5 also shows the autocorrelation of the measured signal as a function of φ. Since the autocorrelation function computes relative delays, it intrinsically separates the echo periods from phase offsets around the ring. The glimmer signature naturally lives in this autocorrelation space; the glimmer mapping is plotted in the figure as a set of blue $\otimes$'s. Half-period echoes, which can be excited by emission in the midplane, are marked as orange $\otimes$'s.

The subset of echoes that are excited is determined by the set of bound orbits that the source intersects. As the source size increases, the φ width of each response blip increases as nearly closed orbits begin to resample the source. When the hotspot is in the midplane, it excites the half-period echoes. As the hotspot moves away from the midplane, the autocorrelation response drifts away from the orange $\otimes$'s. Also, as the hotspot moves from the midplane to low inclination, it intersects fewer bound orbits (e.g., consider the fact that the prograde and retrograde orbits lie entirely within the midplane). Thus, in addition to producing responses with more complicated delay structures, hotspots off the midplane cannot excite the full spectrum of glimmer echoes.

While the echoes described by glimmer are always excited, geodesic self-intersection excites supplemental echoes based on source features. If the majority of the flux in an observation is produced by a single, transient hotspot, then the additional echo structure may provide a means to infer details about the hotspot. In contrast, if the source emission is extended, diffuse, and stochastic, the supplemental echoes will be washed out by the universal glimmer signature.

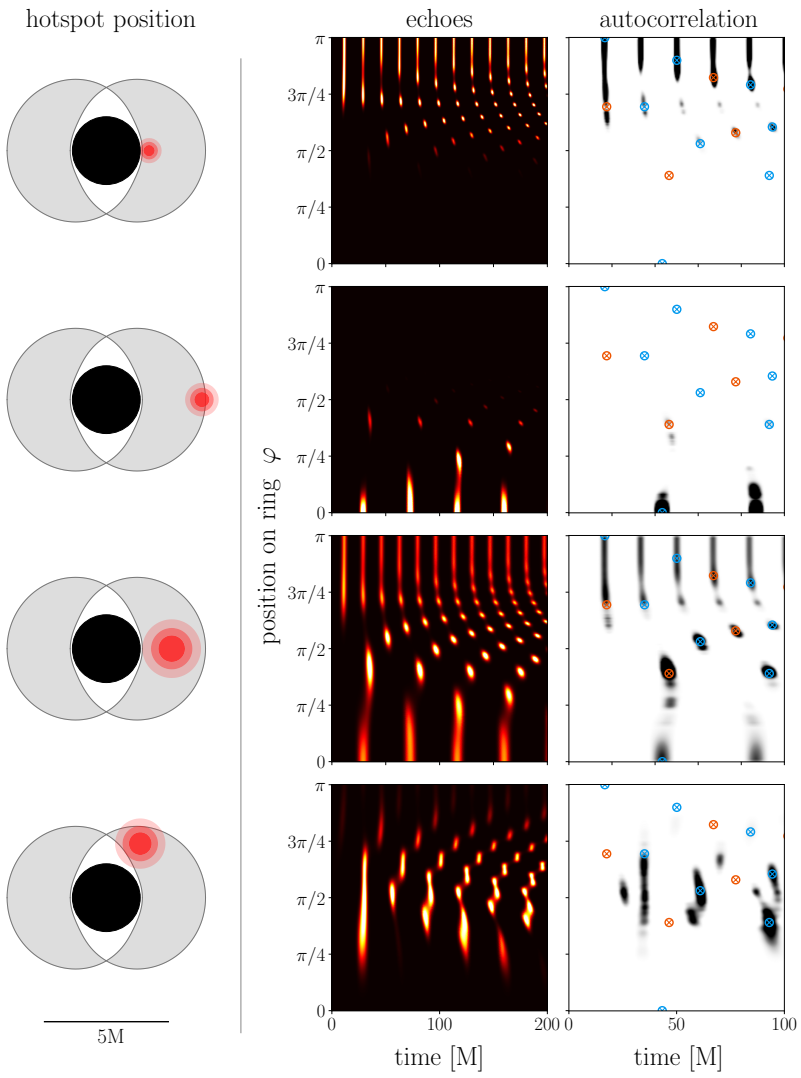


Figure 7.5: Effect of hotspot position and size on echo response. Left column: position and size of hotspot (on Keplerian orbit with $\sigma_t = 1\,\mathrm{M}$) in space around a black hole (black circle) with $a_* = 15/16$. The gray shaded crescents mark the region containing the bound orbits. Central column: echoes produced by the hotspot (as in Figure 7.4). Right column: autocorrelation of echoes. Blue $\otimes$'s denote pure glimmer echoes, and orange $\otimes$'s correspond to midplane echoes. Midplane echoes are excited in the first three rows. When the hotspot is raised above the midplane in the bottom panel, the supplemental echoes no longer peak at $\frac{1}{2}$ the fundamental glimmer period. In contrast, the glimmer echoes are universal; if a given φ exhibits any echoes, it will exhibit echoes at the blue $\otimes$'s. Since the hotspot in the fourth panel does not intersect the (equatorial) prograde and retrograde orbits, it does not excite echoes along those φ. An animation showing how the echoes and autocorrelation change as a function of hotspot size and position is available online.

7.5 Discussion

We have described black hole glimmer, a position- and time-dependent effect in black hole images that is due to bound and closed photon orbits. The set of echo periods and arrival locations that comprise glimmer uniquely encodes the properties of the spacetime and provides a direct way to measure black hole mass and angular momentum. Since the universal glimmer signature is determined only by the underlying geometry, it is separable from astrophysical and plasma uncertainties in the emission model. We now briefly present a simulated glimmer measurement at finite resolution, discuss the generality of the hotspot model, and consider extensions to our analysis.

7.5.1 Measuring glimmer with finite resolution

Recently, Hadar et al. (2021) discussed the potential of performing a measurement of autocorrelations for a nearly top-down black hole with a thin surface emissivity in the midplane, and Chesler et al. (2021) treated the potential of performing a measurement of coherent autocorrelations in a similar context. The clearest glimmer signatures are encoded in the time- and position-dependent intensities along the critical curve, but performing a measurement with sufficient resolution to resolve the detailed spatial dependence may be technologically impractical. Although evaluating the feasibility of a detailed measurement of glimmer is beyond the scope of this paper, we now perform a simple example glimmer measurement from a simulated low-resolution observation.

Even with limited resolution, the echo structure can be probed by comparing the light curves produced by different regions in an image. Figure 7.6 shows the light curves measured in different regions of an angular image decomposition as seen by an observer at $i = 90°$. The different panels show the decomposition for a numerically simulated hotspot ($r_0 = 3\,\mathrm{M}$, $\sigma_r = 0.8\,\mathrm{M}$, and $\sigma_T = 20\,\mathrm{M}$) on a Keplerian orbit in the midplane around three different black holes with spins $a_* = 1/4,\ 15/16$, and $31/32$. After the initial transient decays, the

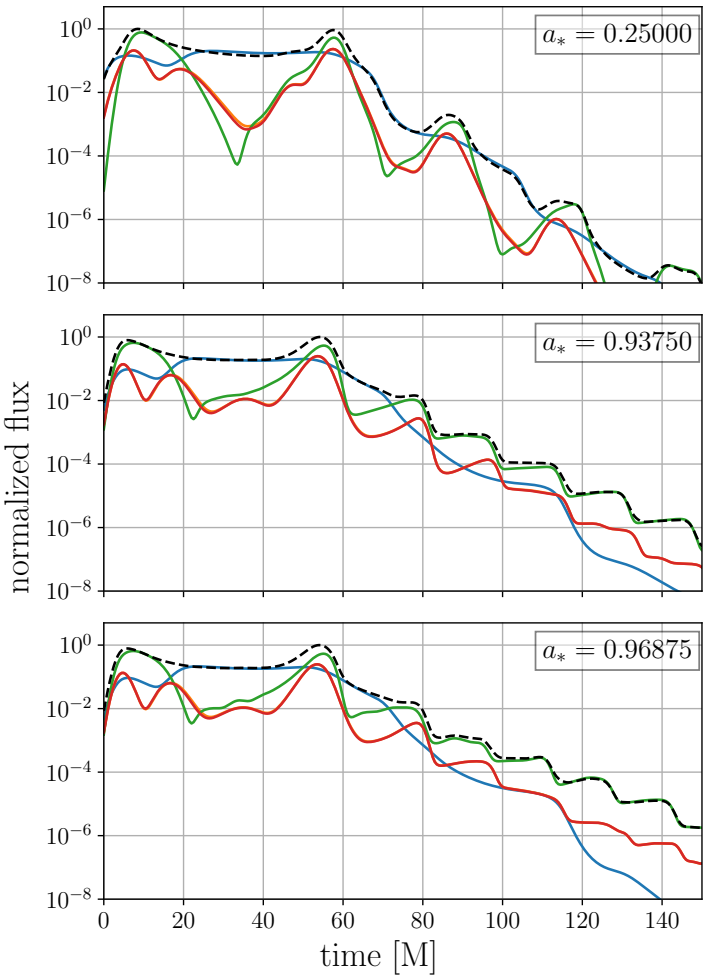


Figure 7.6: Quadrant-based light curve decomposition for a hotspot at $r_0 = 3$ M and $\sigma_r = 0.8$ M with a flat emissivity profile versus frequency orbiting around black holes with $a_* = 1/4, 15/16$, and $31/32$. The full light curve (black dashed line) is divided into four image quadrants: retrograde-centered $\varphi \in (\pi/4, -\pi/4)$ (blue), prograde-centered (green) $\varphi \in (3\pi/4, 5\pi/4)$, and two remaining (red) regions. The prograde orbit echoes are strongest because they have the smallest Lyapunov exponents. The initial transient to $t \approx 60$ GM/c^3 is produced during the time that the hotspot is active and orbiting the hole.

unresolved (total) light-curve signal is dominated by the prograde echo period, but since the different light curves correspond to different parts of the image—and thus cover different arcs along the critical curve—they echo with different periods.

As Figure 7.6 shows, it may be possible to resolve different echo periods among the different light curves. Even if the precise time delay between peaks in a single light curve

is not measurable, the spread in delays could be used to bound the spin and mass of the system. A rough comparison of the delay periods could also be performed to bound the orientation of the curve and infer the orientation of the spin axis of the hole on the sky.

7.5.2 Hotspot model generality

We evaluated the glimmer signature produced by a transient hotspot in a Keplerian orbit around a black hole; however, this toy emission model may not be representative of typical astrophysical scenarios. Real-world emitters likely follow more complicated trajectories, and changing the velocity of the emitting material can contribute to effective source anisotropy via relativistic beaming effects. As the relative amplitudes of the signals emitted in different directions change, the relative amplitude of echo signatures at different φ will change as well. Additionally, real emission sources may comprise diffuse structures that are not well approximated by localized hotspot-like emission.

Nevertheless, our model approximation is not restrictive because any realistic source can be decomposed into a set of localized, transient emission features. In the limit of zero optical depth, the echo signatures produced by different hotspots can be added linearly, with each component weighted by the relative strength of the underlying emission feature. Thus, although the emission structure can influence the relative intensity of the echoes, the position-dependent delay spectrum is universal.

If the emission source is localized, then it may be possible to constrain the vertical location of the emission by subtracting the glimmer echoes from the full observed echo (autocorrelation) spectrum. We note, however, that the transient signatures produced by direct emission have much higher flux relative to the echoes and are therefore likely to provide a much more efficient means of inferring source structure.

It may also be possible to measure the glimmer produced by an emitter far from the hole, e.g., especially in the case of a stellar mass black hole. If a coherent light source is located directly behind the hole relative to the observer, then the light paths between the source

and the observer will undergo the same lensing effects as they pass the hole.

7.5.3 Limitations and extensions

In our treatment, we assumed that the emission was independent of frequency; in sources with nontrivial frequency dependence, redshift effects can change the intensity of light received at different φ around the critical curve and thus may influence the sensitivity required for a measurement. If the emission spectrum has a characteristic frequency, then it may be possible to decompose an unresolved light curve and measure glimmer by comparing the dominant echo periods of different components of the spectrum. Since each bound orbit follows a different latitudinal profile, different orbits can intersect the source at different angles relative to the source motion. Frequency shift is controlled by this angle, so different orbits (and thus different echo periods) will peak with different characteristic frequencies. We also neglected the effect of optical depth, which decreases the strength of high-order echoes compared to the analytic Lyapunov treatment. Since optical depth is a function of wavelength, it may be desirable to observe at frequencies where the optical depth of the plasma is minimal.

More broadly, our treatment neglected the initial transient (which is a strong function of the position, shape, and dynamics of the source). An analysis of general relativistic magnetohydrodynamics simulations may be required to study the detailed complexities of a true observation; an analytic treatment of correlations must faithfully reproduce the spatial structure of the emission source, since source position determines which echoes are excited.

7.6 Summary

We have described black hole glimmer, a position- and time-dependent echoing in black hole images that is due to bound and closed photon orbits. The set of echo periods and arrival locations that comprise glimmer uniquely encodes the properties of the spacetime and

provides a robust, independent probe of black hole mass and angular momentum. Glimmer makes precise predictions that could be used to directly test the Kerr hypothesis. Since glimmer is determined only by the underlying geometry, its signature is separable from astrophysical and plasma uncertainties in the emission model. We have used numerical simulations to demonstrate that the glimmer signature may be observable even in a limited-resolution measurement.

7.7 Kerr equations of motion in detail

The Kerr metric is cyclic in the t and ϕ coordinates, so it has two killing fields, ∂_t and ∂_ϕ. These correspond to the conserved conjugate momenta p_t and p_ϕ that are canonically associated with (negative) energy at infinity and angular momentum about the spin axis of the hole. Carter (1968) identified a third conserved quantity $\mathcal{Q}$ that is associated with a second-order killing tensor field and is physically related to the θ velocity of a particle as it passes through the midplane.

For a test particle with four-momentum p^μ, the three conserved quantities (7.18)–(7.20) along with the particle's mass, $-\mu^2 = p^\mu p_\mu$, uniquely determine the geodesics in the Kerr spacetime:

$$\mathcal{Q} = p_\theta^2 + \cos^2\theta \left(a_*^2\left(\mu^2 - p_t^2\right) + p_\phi^2/\sin^2\theta\right) \tag{7.18}$$

$$\widetilde{E} \equiv -p_t \tag{7.19}$$

$$L_z \equiv p_\phi \tag{7.20}$$

Photons follow null geodesics $x^\mu(\lambda)$, where λ is an affine parameter such that $u^\mu = \dot{x}^\mu \equiv \mathrm{d}x^\mu/\mathrm{d}\lambda$ and $u^\mu u_\mu = 0$. The equations of motion for null geodesics around a black hole are scale invariant with respect to the mass of the hole and can be written as a one-parameter

set of ordinary differential equations (Carter, 1968; Bardeen et al., 1972):

$$\Delta\Sigma\dot{t} = \left(\left(r^2 + a_*^2\right)^2 - \Delta a_*^2 \sin^2\theta\right)\widetilde{E} - 2ra_*L_z \tag{7.21}$$

$$\Sigma^2\dot{r}^2 = \widetilde{E}^2 r^4 + \left(a_*^2E^2 - L_z^2 - \mathcal{Q}\right)r^2 + 2\left(\left(a_*\widetilde{E} - L_z\right)^2 + \mathcal{Q}\right)r - a_*^2\mathcal{Q} \tag{7.22}$$

$$\Sigma^2\dot{\theta}^2 = \mathcal{Q} - \left(L_z^2/\sin^2\theta - \widetilde{E}^2a_*^2\right)\cos^2\theta \tag{7.23}$$

$$\Delta\Sigma\dot{\phi} = 2ra_*\widetilde{E} + (\Sigma - 2r)\,L_z/\sin^2\theta. \tag{7.24}$$

Since photons are massless and their paths are independent of their energies $\widetilde{E}$, it is convenient to normalize both L_z and $\mathcal{Q}$ by $\widetilde{E}$ to define new constants of motion Φ and Q. These two constants can be written in terms of the orbit radius r and are often used to parameterize the bound orbits around a hole of a given spin (see the equivalent Equations 7.6 and 7.7).

$$\Phi \equiv L_z/\widetilde{E} = -\frac{r^3 - 3r^2 + a_*^2r + a_*^2}{a_*\,(r-1)} \tag{7.25}$$

$$Q \equiv \mathcal{Q}/\widetilde{E}^2 = -\frac{r^3\left(r^3 - 6r^2 + 9r - 4a_*^2\right)}{a_*^2\,(r-1)^2}. \tag{7.26}$$

By definition, spherical orbits lie at fixed, unchanging radii, $\dot{r} = \ddot{r} = 0$. By solving Equation 7.22 for these two conditions and rejecting nonphysical solutions (see Teo 2003 for more detail), we reproduce Equation 7.3 and find that spherical orbits must lie between

$$r_\pm = 2\left(1 + \cos\left(\frac{2}{3}\cos^{-1}\pm a_*\right)\right). \tag{7.27}$$

It is convenient to introduce a new variable for the latitudinal coordinate θ. Substituting (7.25) and (7.26) and writing $u \equiv \cos\theta$, Equation 7.23 can be written as a fourth order

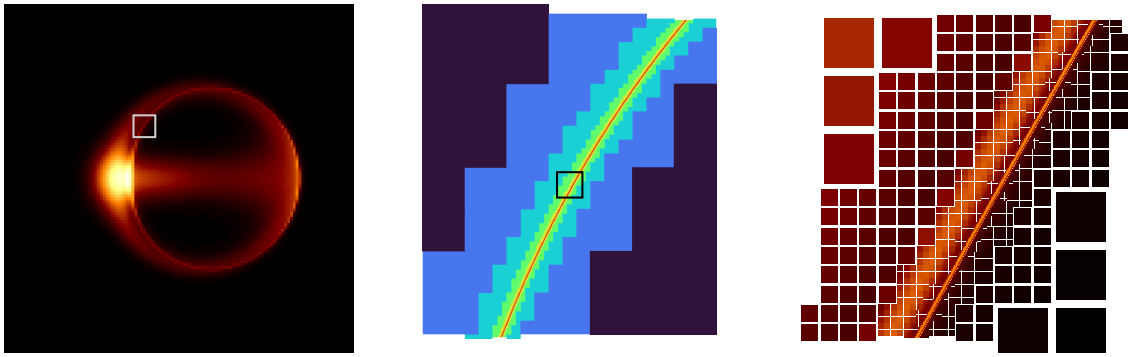


Figure 7.7: Left panel: Composite synthetic black hole image rendered on uniform grid composed by summing all subpixels within lowest-resolution image grid pixels. Center panel: Degree of refinement in region near the critical curve; full panel corresponds to white frame in left panel. Right panel: Pixel-centered intensities shown for pixel at all refinement levels, showing exponentially shrinking self-similar subring structure; full panel area corresponds to black frame in center panel.

polynomial in u:

$$\Sigma^2 \dot{\theta}^2 = Q - \left(\Phi^2 / \sin^2\theta - a_*^2\right) \cos^2\theta \tag{7.28}$$

$$\Sigma^2 \dot{u}^2 = Q - \left(Q + \Phi^2 - a_*^2\right) u^2 - a_*^2 u^4 \tag{7.29}$$

$$= -a_*^2 \left(u^2 - u_+^2\right)\left(u^2 - u_-^2\right). \tag{7.30}$$

The roots of the above expression are given by (see also Equation 7.5)

$$u_\pm^2 \equiv \frac{a_*^2 - Q - \Phi^2 \pm \sqrt{\left(Q + \Phi^2 - a_*^2\right)^2 + 4a_*^2 Q}}{2a_*^2}. \tag{7.31}$$

In terms of u, the two other equations of motion are

$$\dot{t} = \frac{\left(a_*^2 + r^2\right)^2 - a_*^2 \Delta \left(1 - u^2\right) - 2a_* \Phi r}{\Delta \Sigma} \tag{7.32}$$

$$\dot{\phi} = \frac{2a_* r + \Phi \left(r^2 + a_*^2 u^2 - 2r\right)\left(1 - u^2\right)^{-1}}{\Delta \Sigma}. \tag{7.33}$$

To compute the $\Delta\phi$ azimuthal precession identified in §7.2, we integrate $\mathrm{d}\phi/\mathrm{d}u$ over four

quarter cycles,

$$\begin{aligned}\Delta\phi &= 4\int_0^{\theta_+} \frac{\mathrm{d}\phi}{\mathrm{d}\theta}\,\mathrm{d}\theta = 4\int_0^{u_+} \frac{\mathrm{d}\phi}{\mathrm{d}u}\,\mathrm{d}u \\ &= \frac{4}{\sqrt{-u_-^2}}\left(\frac{2r - a_*\Phi}{\Delta}K\left(\frac{u_+^2}{u_-^2}\right) + \frac{\Phi}{a_*}\Pi\left(u_+^2, \frac{u_+^2}{u_-^2}\right)\right). \end{aligned} \tag{7.34}$$

Here, we have written the answer in terms of complete elliptic integrals by expressing $\mathrm{d}u$ in terms of the product written in Equation 7.30. The Δt integral reported in Equation 7.8 is solved in the same way.

For $a_* \neq 0$, the ratio $\Delta\phi/2\pi$ cannot be one; thus, a complete azimuthal cycle will not correspond to a complete latitude cycle. The sign of $\Delta\phi$ corresponds to the net displacement of the orbit (as either prograde or retrograde) and mirrors the sign of Φ. The radius of the polar orbit, which has $\Phi = 0$, is given by (e.g., Teo, 2003)

$$r = 1 + 2\sqrt{1 - \frac{1}{3}a_*^2}\cos\left(\frac{1}{3}\arccos\frac{1 - a_*^2}{\left(1 - \frac{1}{3}a_*^2\right)^{3/2}}\right). \tag{7.35}$$

7.8 Numerical image generation details

The ray traced results presented in this paper were produced using a custom version of the `ipole` code (Mościbrodzka & Gammie, 2018). Most observer-to-emitter codes like `ipole` solve the radiation transport equations along a single geodesic per resolution element (pixel). This approach is reasonable when the difference between neighboring geodesic trajectories is small (such as when the pixels are small or when the geodesics do not pass close to the photon sphere). Since we study the neighborhood of the critical curve, we must resolve the differences between geodesics as they begin to wind around the hole. Thus, in our case, the pixel-centered method with a fixed grid fails, since the widths of the subrings decrease exponentially. We deal with this issue by adaptively concentrating resolution elements near the critical curve.

The modified code first identifies the geodesics that have the longest path lengths relative to their neighbors and then constructs a connected set of pixels (the set containing the identified geodesics) to refine. Each pixel is refined into a 3x3 set of pixels centered around the original geodesic. The process is repeated until a stopping criterion is met. Figure 7.7 shows the image produced by ray tracing on a grid with eight refinement levels, a schematic of refinement levels near the critical curve, and the multiple self-similar subrings produced by each of the imaged turnings.

Chapter 8

Outlook

‘‘ . . . I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.”

Isaac Newton

In his inaugural lecture at the Cavendish Laboratory in 1871, James Clerk Maxwell cautioned against scientific hubris, noting that “the history of science shews that even during that phase of her progress in which she devotes herself to improving the accuracy of the numerical measurement of quantities with which she has long been familiar, she is preparing the materials for the subjugation of new regions.” Indeed, a few years later, Maxwell published his transformative Treatise on Electricity and Magnetism, and the theoretical and experimental advances made in the following decades would significantly reshape our understanding of the physical world. Addressing the slight exceptions to the contemporary “laws of nature” led to the development of entirely new fields like relativity, quantum mechanics and its descendants, and cosmology.

We are now entering an exciting epoch of precision black hole astrophysics. The last decade has witnessed an explosion of new science-enabling measurements from gravitational waves to high resolution imaging and astrometry. These advances let us study black holes with extreme precision down to horizon scales, where strong gravity reigns supreme. Even if we do not uncover exotic new physics, we are still poised to address myriad questions about astrophysical plasma physics, accretion theory, and general relativity. This dissertation has

expatiated on a series of analysis- and theory-based projects that use the first generation of observational data to study black hole accretion apropos of the connection between supermassive black holes and relativistic jets, but the ongoing observational experiments with increased imaging capabilities, improved sensitivity, and support for multi-frequency measurements provide an immediate, clear extension to my work.

In the coming years, theory will play an important role in guiding the next generation of experiments. High resolution measurements of the photon ring impel the establishment of long, space-bound baselines and are highly desired for their ability to both constrain the spin of the black hole and act as a new high-precision test of general relativity at a causal edge of the universe. Uncertainties in the microphysical details of the plasma thermodynamics, especially in the jet region, motivate observations that produce detailed, sensitive measurements of circular polarization. As observations and our understanding of black hole accretion improve, we may develop methods to perform black hole tomography and reconstruct a full three-dimensional map of the plasma near the hole.

Still, there is much work to be done on the analysis front. The parameter-discrimination task is particularly challenging, since a wide variety of black hole accretion models produce similar images. Machine learning methods have been shown to be capable of differentiating between different accretion models (van der Gucht et al., 2020; Lin et al., 2020), although the robustness of the results is yet poorly understood. Efforts to understand the utility of machine learning methods in the direct VLBI data domain are underway, as are tests of the algorithms' abilities to classify images from simulations run with parameters that it had not been exposed to during training.

Other universal parameter inference methods, like measuring black hole glimmer and the shape of the photon ring, are still new and relatively unexplored. Beyond glimmer, theoretical advances have been made in characterizing autocorrelations in Kerr (Hadar et al., 2021; Chesler et al., 2021), but there are no immediate studies of the signatures in theories that deviate from general relativity, and no rigorous explorations of measurement feasibility

have been performed. Gralla et al. (2020) investigated ring shape measurements using toy models of emission, but the signatures produced in the real world from a thick, turbulent flow may differ dramatically from the analytic model. Addressing this question will likely require high resolution fluid simulations as well as high resolution images, perhaps generated using an adaptive ray-tracing code (e.g., Gelles et al. 2021b).

The simulation front remains open as well. Since tilted disks can produce dramatically different image morphologies (see the considerations in Fragile et al. 2007; Dexter & Fragile 2013), it would be interesting to see the results of a full parameter survey that sweeps over tilt in addition to the spin and magnetic flux on the hole. Radiative simulations lie on the horizon as well—algorithmic advances (e.g., Ryan & Dolence 2020) and the proliferation of GPUs as a computational tool suggest that large scale parameter surveys of both SANE and MAD accretion disks, possibly even with tilt, may be feasible in the near future.

The jet composition question is also far from settled. Although drizzle pair production seems to disallow pair cascades in M87, a wide swath of parameter space has yet to be considered. A more general treatment of pair production would account for the drizzle process, perhaps performed *in situ* during a radiation simulation, as well as the pairs produced during cascades. Although cascades cannot be launched from regular ideal GRMHD simulations, a subgrid model for gaps (e.g., Parfrey et al. 2019) based on violations of the Goldreich–Julian conditions on the local charge density could provide a reasonable first attempt on the way to full hybrid kinetic–GRMHD simulations. Future treatment of the pair production problem might also aim to represent the positron–electron pair plasma as a separate fluid.

The future of Michell and Laplace's dark stars is bright.

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, Physical Review Letters, 116, 061102

—. 2016b, Physical Review Letters, 116, 221101

—. 2017, Physical Review Letters, 119, 161101

—. 2019a, The Astrophysical Journal Letters, 882, L24

—. 2019b, Physical Review X, 9, 031040

Abramowicz, M. A., Chen, X., Kato, S., Lasota, J.-P., & Regev, O. 1995, The Astrophysical Journal Letters, 438, L37

Abuter, R., Amorim, A., Anugu, N., et al. 2018, A&A, 615, L15

Agol, E. 2000, The Astrophysical Journal Letters, 538, L121

Akiyama, K., Lu, R.-S., Fish, V. L., et al. 2015, The Astrophysical Journal, 807, 150

Akiyama, K., Kuramochi, K., Ikeda, S., et al. 2017a, ApJ, 838, 1

Akiyama, K., Ikeda, S., Pleau, M., et al. 2017b, Astron. J., 153, 159

Alfvén, H. 1942, Nature, 150, 405

Amarilla, L., & Eiroa, E. F. 2013, Physical Review D, 87, 044057

Amarilla, L., Eiroa, E. F., & Giribet, G. 2010, Physical Review D, 81, 124045

An, T., Goss, W. M., Zhao, J.-H., et al. 2005, The Astrophysical Journal Letters, 634, L49

Anantua, R., Emami, R., Loeb, A., & Chael, A. 2020a, ApJ, 896, 30

Anantua, R., Ressler, S., & Quataert, E. 2020b, MNRAS, 493, 1404

Asada, K., & Nakamura, M. 2012, The Astrophysical Journal Letters, 745, L28

AXFORD, W. I. 1960, The Quarterly Journal of Mechanics and Applied Mathematics, 13, 314

Baade, W., & Minkowski, R. 1954, The Astrophysical Journal, 119, 215

Baganoff, F. K., Maeda, Y., Morris, M., et al. 2003, The Astrophysical Journal, 591, 891

Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214

—. 1998, Reviews of Modern Physics, 70, 1

Balick, B., & Brown, R. L. 1974, The Astrophysical Journal, 194, 265

Ball, D., Özel, F., Psaltis, D., Chan, C.-K., & Sironi, L. 2018, ApJ, 853, 184

Ballantyne, D. R., Özel, F., & Psaltis, D. 2007, The Astrophysical Journal Letters, 663, L17

Balsara, D. S., & Spicer, D. S. 1999, Journal of Computational Physics, 149, 270

Bambi, C., & Freese, K. 2009, Physical Review D, 79, 043002

Bardeen, J. M. 1973, Proc. Ecole É Phys. Théorique, 215

Bardeen, J. M., & Petterson, J. A. 1975, The Astrophysical Journal Letters, 195, L65

Bardeen, J. M., Press, W. H., & Teukolsky, S. A. 1972, The Astrophysical Journal, 178, 347

Beckers, J. M. 1993, Annu. Rev. Astron. Astrophys., 31, 13

Beckert, T., & Falcke, H. 2002, Astronomy and Astrophysics, 388, 1106

Berezhko, E. G., & Krymskii, G. F. 1981, Sov. Astron. Lett., 7, 352

Berti, E., Cardoso, V., Gonzalez, J. A., et al. 2007, Physical Review D, 76, 064034

Berti, E., Cardoso, V., & Will, C. M. 2006, Physical Review D, 73, 064030

Beskin, V. S., Istomin, Y. N., & Parev, V. I. 1992, Sov. Astron., 36, 642

Bird, S., Harris, W. E., Blakeslee, J. P., & Flynn, C. 2010, A&A, 524, A71

Bisnovatyi-Kogan, G. S., & Ruzmaikin, A. A. 1974, Astrophysics and Space Science, 28, 45

Blakeslee, J. P., Jordán, A., Mei, S., et al. 2009, ApJ, 694, 556

Blandford, R., Meier, D., & Readhead, A. 2019, Annual Review of Astronomy and Astrophysics, 57, 467

Blandford, R. D., & Payne, D. G. 1982, MNRAS, 199, 883

Blandford, R. D., & Znajek, R. L. 1977, MNRAS, 179, 433

Blank, M., Morris, M. R., Frank, A., Carroll-Nellenback, J. J., & Duschl, W. J. 2016, Monthly Notices of the Royal Astronomical Society, 459, 1721

Boccardi, B., Krichbaum, T. P., Ros, E., & Zensus, J. A. 2017, Astronomy and Astrophysics Review, 25, 4

Bolton, C. T. 1972, Nature Physical Science, 240, 124

Boris, J. P. 1990, in Whither Turbulence? Turbulence at the Crossroads, Vol. 357, 344–353

Boris, J. P., Grinstein, F. F., Oran, E. S., & Kolbe, R. L. 1992, Fluid Dynamics Research, 10, 199

Bower, G. C., Wright, M. C. H., Falcke, H., & Backer, D. C. 2003, The Astrophysical Journal, 588, 331

Bower, G. C., Markoff, S., Dexter, J., et al. 2015, The Astrophysical Journal, 802, 69

Bowyer, S., Byram, E. T., Chubb, T. A., & Friedman, H. 1965, Science, 147, 394

Boyer, R. H., & Lindquist, R. W. 1967, Journal of Mathematical Physics, 8, 265

Breit, G., & Wheeler, J. A. 1934, Phys. Rev., 46, 1087

Britzen, S., Fendt, C., Eckart, A., & Karas, V. 2017, A&A, 601, A52

Broderick, A., & Blandford, R. 2004, Monthly Notices of the Royal Astronomical Society, 349, 994

Broderick, A. E., & Loeb, A. 2005, Monthly Notices of the Royal Astronomical Society, 363, 353

—. 2006, The Astrophysical Journal Letters, 636, L109

—. 2009, Astrophys. J., 697, 1164

Broderick, A. E., Narayan, R., Kormendy, J., et al. 2015, ApJ, 805, 179

Broderick, A. E., & Tchekhovskoy, A. 2015, Astrophys. J., 809, 97

Broderick, A. E., Gold, R., Karami, M., et al. 2020, ApJ, 897, 139

Bromley, B. C., Melia, F., & Liu, S. 2001, The Astrophysical Journal Letters, 555, L83

Bronzwaer, T., Davelaar, J., Younsi, Z., et al. 2018, Astronomy and Astrophysics, 613, A2

Brown, R. L. 1982, The Astrophysical Journal, 262, 110

Buonanno, A., Cook, G. B., & Pretorius, F. 2007, Physical Review D, 75, 124018

Burn, B. J. 1966, Monthly Notices of the Royal Astronomical Society, 133, 67

Canfield, E., Howard, W. M., & Liang, E. P. 1987, The Astrophysical Journal, 323, 565

Cantiello, M., Blakeslee, J. P., Ferrarese, L., et al. 2018, ApJ, 856, 126

Carter, B. 1968, Physical Review, 174, 1559

—. 1979, Gen. Relativ. Einstein Centen. Surv., 294

Chael, A., Narayan, R., & Johnson, M. D. 2019, Monthly Notices of the Royal Astronomical Society, 486, 2873

Chael, A., Rowan, M., Narayan, R., Johnson, M., & Sironi, L. 2018a, Monthly Notices of the Royal Astronomical Society, 478, 5209

Chael, A. A., Johnson, M. D., Bouman, K. L., et al. 2018b, Astrophys. J., 857, 23

Chael, A. A., Johnson, M. D., Narayan, R., et al. 2016, The Astrophysical Journal, 829, 11

Chael, A. A., Narayan, R., & Saḍowski, A. 2017, Mon. Not. R. Astron. Soc., 470, 2367

Chan, C.-k., Medeiros, L., Özel, F., & Psaltis, D. 2018, The Astrophysical Journal, 867, 59

Chan, C.-k., Psaltis, D., & Özel, F. 2013, The Astrophysical Journal, 777, 13

Chandra, M., Gammie, C. F., Foucart, F., & Quataert, E. 2015, ApJ, 810, 162

Chatterjee, K., Liska, M., Tchekhovskoy, A., & Markoff, S. B. 2019, Monthly Notices of the Royal Astronomical Society, 490, 2200

Chatterjee, K., Younsi, Z., Liska, M., et al. 2020, Monthly Notices of the Royal Astronomical Society, 499, 362

Chen, A. Y., & Yuan, Y. 2020, Astrophys. J., 895, 121

Chen, A. Y., Yuan, Y., & Yang, H. 2018, Astrophys. J., 863, L31

Chen, Z., Gou, L., McClintock, J. E., et al. 2016, The Astrophysical Journal, 825, 45

Chesler, P. M., Blackburn, L., Doeleman, S. S., et al. 2021, Class. Quantum Grav., 38, 125006

Clark, B. G. 1980, Astron. Astrophys., 89, 377

Claudel, C.-M., Virbhadra, K. S., & Ellis, G. F. R. 2001, Journal of Mathematical Physics, 42, 818

Cotera, A., Morris, M., Ghez, A. M., et al. 1999, ASP Conf. Ser., 186, 240

Crinquand, B., Cerutti, B., Philippov, A., Parfrey, K., & Dubus, G. 2020, Phys. Rev. Lett., 124, 145101

Cuadra, J., Nayakshin, S., Springel, V., & Di Matteo, T. 2006, Monthly Notices of the Royal Astronomical Society, 366, 358

Curtis, H. D. 1918, Publications of Lick Observatory, 13, 9

Dagum, L., & Menon, R. 1998, IEEE Comput. Sci. Eng., 5, 46

Darwin, C. 1959, Proceedings of the Royal Society of London Series A, 249, 180

Davelaar, J., Mościbrodzka, M., Bronzwaer, T., & Falcke, H. 2018, A&A, 612, A34

Davelaar, J., Olivares, H., Porth, O., et al. 2019, A&A, 632, A2

Davis, S. W., & Tchekhovskoy, A. 2020, Annu. Rev. Astron. Astrophys. Vol 58 P407-439, 58, 407

de Gasperin, F., Orrú, E., Murgia, M., et al. 2012, Astronomy and Astrophysics, 547, A56

De Luca, V., Desjacques, V., Franciolini, G., Malhotra, A., & Riotto, A. 2019, J. Cosmol. Astropart. Phys. Issue 05 Artic. Id 018 2019, 2019, 018

De Villiers, J.-P., & Hawley, J. F. 2003, The Astrophysical Journal, 592, 1060

De Villiers, J.-P., Hawley, J. F., & Krolik, J. H. 2003, The Astrophysical Journal, 599, 1238

De Villiers, J.-P., Hawley, J. F., Krolik, J. H., & Hirose, S. 2005, The Astrophysical Journal, 620, 878

Del Zanna, L., Papini, E., Landi, S., Bugli, M., & Bucciantini, N. 2016, Monthly Notices of the Royal Astronomical Society, 460, 3753

Del Zanna, L., Zanotti, O., Bucciantini, N., & Londrillo, P. 2007, Astron. Astrophys., 473, 11

Dexter, J. 2016, Mon. Not. R. Astron. Soc., 462, 115

Dexter, J., & Agol, E. 2009, The Astrophysical Journal, 696, 1616

Dexter, J., Agol, E., Fragile, P. C., & McKinney, J. C. 2010, ApJ, 717, 1092

Dexter, J., & Fragile, P. C. 2013, Monthly Notices of the Royal Astronomical Society, 432, 2252

Dexter, J., McKinney, J. C., & Agol, E. 2012, Monthly Notices of the Royal Astronomical Society, 421, 1517

Dexter, J., Jiménez-Rosales, A., Ressler, S. M., et al. 2020, Monthly Notices of the Royal Astronomical Society, 494, 4168

Di Matteo, T., Allen, S. W., Fabian, A. C., Wilson, A. S., & Young, A. J. 2003, The Astrophysical Journal, 582, 133

Dibi, S., Drappeau, S., Fragile, P. C., Markoff, S., & Dexter, J. 2012, Monthly Notices of the Royal Astronomical Society, 426, 1928

Dionysopoulou, K., Alic, D., Palenzuela, C., Rezzolla, L., & Giacomazzo, B. 2013, Phys. Rev. D, 88, 044020

Dodds-Eden, K., Porquet, D., Trap, G., et al. 2009, The Astrophysical Journal, 698, 676

Doeleman, S., Agol, E., Backer, D., et al. 2009, Astro2010 Astron. Astrophys. Decadal Surv., 2010, 68

Doeleman, S. S., Weintroub, J., Rogers, A. E. E., et al. 2008, Nature, 455, 78

Doeleman, S. S., Fish, V. L., Schenck, D. E., et al. 2012, Science, 338, 355

Dolence, J. C., Gammie, C. F., Mościbrodzka, M., & Leung, P. K. 2009, The Astrophysical Journal Supplement Series, 184, 387

Dolence, J. C., Gammie, C. F., Shiokawa, H., & Noble, S. C. 2012, The Astrophysical Journal Letters, 746, L10

Eckart, A., & Genzel, R. 1997, Monthly Notices of the Royal Astronomical Society, 284, 576

Einstein, A. 1905, Annalen der Physik, 322, 891

—. 1916, Annalen der Physik, 354, 769

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019a, ApJL, 875, L1

—. 2019b, ApJL, 875, L2

—. 2019c, ApJL, 875, L3

—. 2019d, ApJL, 875, L4

—. 2019e, ApJL, 875, L5

—. 2019f, ApJL, 875, L6

Event Horizon Telescope Collaboration, Akiyama, K., Algaba, J. C., et al. 2021a, ApJL, 910, L12

—. 2021b, ApJL, 910, L13

Falcke, H., & Biermann, P. L. 1995, Astron. Astrophys., 293, 665

Falcke, H., Goss, W. M., Matsuo, H., et al. 1998, The Astrophysical Journal, 499, 731

Falcke, H., Melia, F., & Agol, E. 2000, The Astrophysical Journal Letters, 528, L13

Fejer, J. A. 1964, The Physics of Fluids, 7, 499

Finkelstein, D. 1958, Physical Review, 110, 965

Fish, V. L., Doeleman, S. S., Beaudoin, C., et al. 2011, The Astrophysical Journal Letters, 727, L36

Fishbone, L. G., & Moncrief, V. 1976, The Astrophysical Journal, 207, 962

Ford, A. L., Keenan, B. D., & Medvedev, M. V. 2018, Phys. Rev. Vol. 98 Issue 6 Id063016, 98, 063016

Forum, M. P. 1994, MPI: A Message-Passing Interface Standard, Technical Report, University of Tennessee, USA

Foucart, F., Chandra, M., Gammie, C. F., Quataert, E., & Tchekhovskoy, A. 2017, Monthly Notices of the Royal Astronomical Society, 470, 2240

Fragile, P. C., & Blaes, O. M. 2008, The Astrophysical Journal, 687, 757

Fragile, P. C., Blaes, O. M., Anninos, P., & Salmonson, J. D. 2007, The Astrophysical Journal, 668, 417

Fragile, P. C., & Meier, D. L. 2009, The Astrophysical Journal, 693, 771

Fromm, C. M., Younsi, Z., Baczko, A., et al. 2019, A&A, 629, A4

Gammie, C. F., & Leung, P. K. 2012, The Astrophysical Journal, 752, 123

Gammie, C. F., McKinney, J. C., & Tóth, G. 2003, ApJ, 589, 444

Garofalo, D. 2009, The Astrophysical Journal Letters, 699, L52

Garofalo, D., Evans, D. A., & Sambruna, R. M. 2010, Monthly Notices of the Royal Astronomical Society, 406, 975

Gebhardt, K., Adams, J., Richstone, D., et al. 2011, The Astrophysical Journal, 729, 119

Gelles, Z., Himwich, E., Palumbo, D. C. M., & Johnson, M. D. 2021a, arXiv e-prints, 2105, arXiv:2105.09440

Gelles, Z., Prather, B. S., Palumbo, D. C. M., et al. 2021b, Astrophys. J., 912, 39

Genzel, R., & Eckart, A. 1999, ASP Conf. Ser., 186, 3

Genzel, R., Schödel, R., Ott, T., et al. 2003, Nature, 425, 934

Genzel, R., Thatte, N., Krabbe, A., Kroker, H., & Tacconi-Garman, L. E. 1996, The Astrophysical Journal, 472, 153

Ghez, A. M., Klein, B. L., Morris, M., & Becklin, E. E. 1998, ApJ, 509, 678

Ghez, A. M., Wright, S. A., Matthews, K., et al. 2004, The Astrophysical Journal Letters, 601, L159

Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, The Astrophysical Journal, 689, 1044

Gillessen, S., Genzel, R., Fritz, T. K., et al. 2012, Nature, 481, 51

Gillessen, S., Plewa, P. M., Eisenhauer, F., et al. 2017, The Astrophysical Journal, 837, 30

Goddi, C., Martí-Vidal, I., Messias, H., et al. 2021, ApJL, 910, L14

Gold, R., Broderick, A. E., Younsi, Z., et al. 2020, ApJ, 897, 148

Goldreich, P., & Julian, W. H. 1969, Astrophys. J., 157, 869

Goldston, J. E., Quataert, E., & Igumenshchev, I. V. 2005, The Astrophysical Journal, 621, 785

Goldwurm, A., Cordier, B., Paul, J., et al. 1994, Nature, 371, 589

Gralla, S. E., & Lupsasca, A. 2020, Physical Review D, 101, 044031

Gralla, S. E., Lupsasca, A., & Marrone, D. P. 2020, Phys. Rev. Vol. 102 Issue 12 Artic. Id124004, 102, 124004

Gravity Collaboration. 2017, The Messenger, 170, 10

Gravity Collaboration, Abuter, R., Amorim, A., et al. 2018, Astronomy and Astrophysics, 618, L10

—. 2020a, Astron. Astrophys., 636, L5

Gravity Collaboration, Jiménez-Rosales, A., Dexter, J., et al. 2020b, Astronomy and Astrophysics, 643, A56

Greene, J. E., & Ho, L. C. 2007, The Astrophysical Journal, 670, 92

Greenstein, J. L. 1963, Nature, 197, 1041

Greenstein, J. L., & Matthews, T. A. 1963, The Astronomical Journal, 68, 279

Greenstein, J. L., & Schmidt, M. 1964, The Astrophysical Journal, 140, 1

Hada, K., Doi, A., Kino, M., et al. 2011, Nature, 477, 185

Hada, K., Kino, M., Doi, A., et al. 2013, ApJ, 775, 70

—. 2016, ApJ, 817, 131

Hada, K., Park, J. H., Kino, M., et al. 2017, Publications of the Astronomical Society of Japan, 69, 71

Hadar, S., Johnson, M. D., Lupsasca, A., & Wong, G. N. 2021, Phys. Rev. D, 103, 104038

Haller, J. W. 1992, PhD thesis, University of Arizona

Hamaker, J. P., & Bregman, J. D. 1996, Astronomy and Astrophysics Supplement Series, 117, 161

Hanawa, T. 1989, The Astrophysical Journal, 341, 948

Harms, R. J., Ford, H. C., Tsvetanov, Z. I., et al. 1994, The Astrophysical Journal Letters, 435, L35

Harris, D. E., Cheung, C. C., Stawarz, Ł., Biretta, J. A., & Perlman, E. S. 2009, The Astrophysical Journal, 699, 305

Hawley, J. F., Gammie, C. F., & Balbus, S. A. 1995, The Astrophysical Journal, 440, 742

Hawley, J. F., & Krolik, J. H. 2006, The Astrophysical Journal, 641, 103

Hazard, C., Mackey, M. B., & Shimmins, A. J. 1963, Nature, 197, 1037

Heinz, S., & Begelman, M. C. 1997, The Astrophysical Journal, 490, 653

Hilbert, D. 1917, Nachrichten Von Ges. Wiss. Zu Gött. Math.-Phys. Kl., 53

Himwich, E., Johnson, M. D., Lupsasca, A., & Strominger, A. 2020, Physical Review D, 101, 084020

Hioki, K., & Maeda, K.-I. 2009, Physical Review D, 80, 024042

Hirose, S., Krolik, J. H., De Villiers, J.-P., & Hawley, J. F. 2004, The Astrophysical Journal, 606, 1083

Hirotani, K. 1989, Ultra-Hot Plasmas Electron-Positron Pairs Astrophys., 6, 1

Hirotani, K., & Okamoto, I. 1998, Astrophys. J. Vol. 497 Issue 2 Pp 563-572, 497, 563

Hirotani, K., & Pu, H.-Y. 2016, Astrophys. J., 818, 50

Hjellming, R. M., & Rupen, M. P. 1995, Nature, 375, 464

Hobbs, G., Archibald, A., Arzoumanian, Z., et al. 2010, Classical and Quantum Gravity, 27, 084013

Högbom, J. A. 1974, Astron. Astrophys. Suppl. Ser., 15, 417

Howes, G. G. 2010, Monthly Notices of the Royal Astronomical Society, 409, L104

Hoyle, F. 1966, Nature, 209, 751

Hoyle, F., & Burbidge, G. R. 1966, The Astrophysical Journal, 144, 534

Hulse, R. A., & Taylor, J. H. 1975, The Astrophysical Journal Letters, 195, L51

Huygens, C. 1690, Traite de la lumiere. Où sont expliquées les causes de ce qui luy arrive dans la reflexion, & dans la refraction. (chez Pierre Vander Aa marchand libraire)

Ichimaru, S. 1977, The Astrophysical Journal, 214, 840

Igumenshchev, I. V., Narayan, R., & Abramowicz, M. A. 2003, The Astrophysical Journal, 592, 1042

Israel, W. 1967, Physical Review, 164, 1776

—. 1968, Communications in Mathematical Physics, 8, 245

Jiang, G.-S., & Shu, C.-W. 1996, Journal of Computational Physics, 126, 202

Jiménez-Rosales, A., & Dexter, J. 2018, Monthly Notices of the Royal Astronomical Society, 478, 1875

Johannsen, T. 2013, The Astrophysical Journal, 777, 170

Johannsen, T., & Psaltis, D. 2010, The Astrophysical Journal, 718, 446

Johnson, B. M., & Quataert, E. 2007, ApJ, 660, 1273

Johnson, M. D., Lupsasca, A., Strominger, A., et al. 2020, Science Advances, 6, eaaz1310

Jokipii, J. R., & Morfill, G. E. 1990, Astrophys. J., 356, 255

Jones, T. W., & Hardee, P. E. 1979, The Astrophysical Journal, 228, 268

Junor, W., Biretta, J. A., & Livio, M. 1999, Nature, 401, 891

Jüttner, F. 1911, Annalen der Physik, 339, 856

Kahn, H. 1950, Nucleonics, 6, 60

Kamionkowski, M., & Kovetz, E. D. 2016, Annual Review of Astronomy and Astrophysics, 54, 227

Kato, S. 2001, Publications of the Astronomical Society of Japan, 53, L37

Kawazura, Y., Barnes, M., & Schekochihin, A. A. 2019, Proceedings of the National Academy of Science, 116, 771

Kellermann, K. I., & Moran, J. M. 2001, Annu. Rev. Astron. Astrophys., 39, 457

Kerr, R. P. 1963, Physical Review Letters, 11, 237

Kim, J.-Y., Krichbaum, T. P., Lu, R.-S., et al. 2018, Astronomy and Astrophysics, 616, A188

Kimura, S. S., & Toma, K. 2020, Astrophys. J. Vol. 905 Issue 2 Id178, 905, 178

Kino, M., Takahara, F., Hada, K., et al. 2015, ApJ, 803, 30

Kino, M., Takahara, F., Hada, K., & Doi, A. 2014, ApJ, 786, 5

Koide, S., Shibata, K., & Kudoh, T. 1999, The Astrophysical Journal, 522, 727

Kojima, Y. 1991, Monthly Notices of the Royal Astronomical Society, 250, 629

Komissarov, S., & Porth, O. 2021, New Astron. Rev. Vol. 92 Artic. Id 101610, 92, 101610

Komissarov, S. S. 1999, Monthly Notices of the Royal Astronomical Society, 303, 343

Kormendy, J., & Richstone, D. 1995, Annual Review of Astronomy and Astrophysics, 33, 581

Kovalev, Y. Y., Lister, M. L., Homan, D. C., & Kellermann, K. I. 2007, The Astrophysical Journal Letters, 668, L27

Krolik, J. H. 1999, Act. Galact. Nucl. Cent. Black Hole Galact. Environ. Julian H Krolik Princet.

Kunz, M. W., Stone, J. M., & Quataert, E. 2016, Phys. Rev. Lett., 117, 235101

Kuo, C. Y., Asada, K., Rao, R., et al. 2014, The Astrophysical Journal Letters, 783, L33

Kuramochi, K., Akiyama, K., Ikeda, S., et al. 2018, The Astrophysical Journal, 858, 56

Lacy, J. H., Townes, C. H., Geballe, T. R., & Hollenbach, D. J. 1980, The Astrophysical Journal, 241, 132

Lacy, J. H., Townes, C. H., & Hollenbach, D. J. 1982, The Astrophysical Journal, 262, 120

Landi Degl'Innocenti, E., & Landi Degl'Innocenti, M. 1985, Sol. Phys., 97, 239

Laor, A. 1991, The Astrophysical Journal, 376, 90

Laurent, P., & Titarchuk, L. 2018, ApJ, 859, 89

Lemaître, G. 1933, Annales de la Société Scientifique de Bruxelles, 53

Lense, J., & Thirring, H. 1918, Physikalische Zeitschrift, 19

Leung, P. K., Gammie, C. F., & Noble, S. C. 2011, The Astrophysical Journal, 737, 21

Levermore, C. D. 1984, Journal of Quantitative Spectroscopy and Radiative Transfer, 31, 149

Levinson, A., & Cerutti, B. 2018, Astron. Astrophys., 616, doi:10.1051/0004-6361/201832915

Levinson, A., & Rieger, F. 2011, Astrophys. J. Vol. 730 Issue 2 Artic. Id 123, 730, 123

Li, Y.-P., Yuan, F., & Xie, F.-G. 2016, ApJ, 830, 78

Li, Y.-R., Yuan, Y.-F., Wang, J.-M., Wang, J.-C., & Zhang, S. 2009, ApJ, 699, 513

Lin, J. Y.-Y., Wong, G. N., Prather, B. S., & Gammie, C. F. 2020, in ML Interpretability for Scientific Discovery

Liska, M., Hesp, C., Tchekhovskoy, A., et al. 2018, Monthly Notices of the Royal Astronomical Society, 474, L81

Liu, H. B., Wright, M. C. H., Zhao, J.-H., et al. 2016a, Astronomy and Astrophysics, 593, A44

—. 2016b, Astronomy and Astrophysics, 593, A107

Loureiro, N. F., Schekochihin, A. A., & Cowley, S. C. 2007, Phys. Plasmas, 14, 100703

Luminet, J.-P. 1979, Astronomy and Astrophysics, 75, 228

Ly, C., Walker, R. C., & Junor, W. 2007, The Astrophysical Journal, 660, 200

Lynden-Bell, D. 1969, Nature, 223, 690

—. 2006, Monthly Notices of the Royal Astronomical Society, 369, 1167

Lynden-Bell, D., & Rees, M. J. 1971, Monthly Notices of the Royal Astronomical Society, 152, 461

Mahadevan, R., & Quataert, E. 1997, ApJ, 490, 605

Mao, S. A., Dexter, J., & Quataert, E. 2017, Mon. Not. R. Astron. Soc., 466, 4307

Marrone, D. P., Moran, J. M., Zhao, J.-H., & Rao, R. 2006, The Astrophysical Journal, 640, 308

—. 2007, The Astrophysical Journal Letters, 654, L57

Marshall, M. D., Avara, M. J., & McKinney, J. C. 2018, Monthly Notices of the Royal Astronomical Society, 478, 1837

Matthews, T. A., & Sandage, A. R. 1963, The Astrophysical Journal, 138, 30

Maxwell, J. C. 1865, Philosophical Transactions of the Royal Society of London, 155, 459

McKinney, J. C. 2006, Monthly Notices of the Royal Astronomical Society, 368, 1561

McKinney, J. C., & Blandford, R. D. 2009, Monthly Notices of the Royal Astronomical Society, 394, L126

McKinney, J. C., & Gammie, C. F. 2004, The Astrophysical Journal, 611, 977

McKinney, J. C., & Narayan, R. 2007, Monthly Notices of the Royal Astronomical Society, 375, 513

McKinney, J. C., Tchekhovskoy, A., & Blandford, R. D. 2012, Monthly Notices of the Royal Astronomical Society, 423, 3083

—. 2013, Science, 339, 49

McKinney, J. C., Tchekhovskoy, A., Sadowski, A., & Narayan, R. 2014, Monthly Notices of the Royal Astronomical Society, 441, 3177

Medeiros, L., Psaltis, D., & Özel, F. 2020, The Astrophysical Journal, 896, 7

Merck, M., Bertsch, D. L., Dingus, B. L., et al. 1996, Astronomy and Astrophysics Supplement Series, 120, 465

Mertens, F., Lobanov, A. P., Walker, R. C., & Hardee, P. E. 2016, Astronomy and Astrophysics, 595, A54

Messier, C. 1781, Connoissance des Temps ou des Mouvements Célestes, for 1784, p. 227-267, 227

Mignone, A., Bodo, G., Massaglia, S., et al. 2007, Astrophys. J. Suppl. Ser. Vol. 170 Issue 1 Pp 228-242, 170, 228

Mikhailov, A. G., Piotrovich, M. Y., Gnedin, Y. N., Natsvlishvili, T. M., & Buliga, S. D. 2018, Monthly Notices of the Royal Astronomical Society, 476, 4872

Miller, J. M. 2007, Annual Review of Astronomy and Astrophysics, 45, 441

Miller-Jones, J. C. A., Bahramian, A., Orosz, J. A., et al. 2021, Science, 371, 1046

Miura, A., & Pritchett, P. L. 1982, J. Geophys. Res. Space Phys., 87, 7431

Miyoshi, M., Moran, J., Herrnstein, J., et al. 1995, Nature, 373, 127

Mizuno, Y., Younsi, Z., Fromm, C. M., et al. 2018, Nature Astronomy, 2, 585

Montero-Castaño, M., Herrnstein, R. M., & Ho, P. T. P. 2009, The Astrophysical Journal, 695, 1477

Morales Teixeira, D., Fragile, P. C., Zhuravlev, V. V., & Ivanov, P. B. 2014, The Astrophysical Journal, 796, 103

Moriyama, K., & Mineshige, S. 2015, Publications of the Astronomical Society of Japan, 67, 106

Moriyama, K., Mineshige, S., Honma, M., & Akiyama, K. 2019, The Astrophysical Journal, 887, 227

Morningstar, W. R., Miller, J. M., Reis, R. C., & Ebisawa, K. 2014, The Astrophysical Journal Letters, 784, L18

Mościbrodzka, M. 2020, Monthly Notices of the Royal Astronomical Society, 491, 4807

Mościbrodzka, M., Dexter, J., Davelaar, J., & Falcke, H. 2017, Monthly Notices of the Royal Astronomical Society, 468, 2214

Mościbrodzka, M., & Falcke, H. 2013, Astronomy and Astrophysics, 559, L3

Mościbrodzka, M., Falcke, H., & Shiokawa, H. 2016, Astronomy and Astrophysics, 586, A38

Mościbrodzka, M., Falcke, H., Shiokawa, H., & Gammie, C. F. 2014, Astronomy and Astrophysics, 570, A7

Mościbrodzka, M., & Gammie, C. F. 2018, Monthly Notices of the Royal Astronomical Society, 475, 43

Mościbrodzka, M., Gammie, C. F., Dolence, J. C., & Shiokawa, H. 2011, ApJ, 735, 9

Mościbrodzka, M., Gammie, C. F., Dolence, J. C., Shiokawa, H., & Leung, P. K. 2009, The Astrophysical Journal, 706, 497

Nakamura, M., Asada, K., Hada, K., et al. 2018, ApJ, 868, 146

Narayan, R., Igumenshchev, I. V., & Abramowicz, M. A. 2003, Publications of the Astronomical Society of Japan, 55, L69

Narayan, R., Mahadevan, R., Grindlay, J. E., Popham, R. G., & Gammie, C. 1998, The Astrophysical Journal, 492, 554

Narayan, R., Sädowski, A., Penna, R. F., & Kulkarni, A. K. 2012, Monthly Notices of the Royal Astronomical Society, 426, 3241

Narayan, R., & Yi, I. 1994, The Astrophysical Journal Letters, 428, L13

—. 1995a, The Astrophysical Journal, 444, 231

—. 1995b, The Astrophysical Journal, 452, 710

Narayan, R., Yi, I., & Mahadevan, R. 1995, Nature, 374, 623

Narayan, R., Palumbo, D. C. M., Johnson, M. D., et al. 2021, ApJ, 912, 35

Natarajan, P., & Armitage, P. J. 1999, Monthly Notices of the Royal Astronomical Society, 309, 961

Nathanail, A., Fromm, C. M., Porth, O., et al. 2020, Mon. Not. R. Astron. Soc., 495, 1549

Neilsen, J., Nowak, M. A., Gammie, C., et al. 2013, The Astrophysical Journal, 774, 42

Newman, E. T., Couch, E., Chinnapared, K., et al. 1965, Journal of Mathematical Physics, 6, 918

Noble, S. C., Gammie, C. F., McKinney, J. C., & Del Zanna, L. 2006, The Astrophysical Journal, 641, 626

Noble, S. C., Krolik, J. H., & Hawley, J. F. 2009, The Astrophysical Journal, 692, 411

Noble, S. C., Leung, P. K., Gammie, C. F., & Book, L. G. 2007, Classical and Quantum Gravity, 24, S259

Novikov, I. D., & Thorne, K. S. 1973, Black Holes Astres Occlus, 343

Ogihara, T., Ogawa, T., & Toma, K. 2021, Astrophys. J., 911, 34

Ogilvie, G. I. 2016, J. Plasma Phys., 82, 205820301

Ohanian, H. C. 1987, American Journal of Physics, 55, 428

Oke, J. B. 1963, Nature, 197, 1040

Olivares, H., Younsi, Z., Fromm, C. M., et al. 2020, Monthly Notices of the Royal Astronomical Society, 497, 521

Osmanov, Z., Mignone, A., Massaglia, S., Bodo, G., & Ferrari, A. 2008, A&A, 490, 493

Ostrowski, M. 1990, Astron. Astrophys., 238, 435

—. 1999, Plasma Turbul. Energ. Part. Astrophys., 256

Owen, F. N., Eilek, J. A., & Kassim, N. E. 2000, The Astrophysical Journal, 543, 611

Özel, F., Psaltis, D., & Narayan, R. 2000, Astrophys. J. Vol. 541 Issue 1 Pp 234-249, 541, 234

Palenzuela, C., Lehner, L., Reula, O., & Rezzolla, L. 2009, Monthly Notices of the Royal Astronomical Society, 394, 1727

Palmer, H. P., Rowson, B., Anderson, B., Donaldson, W., & Miley, G. K. 1967, Nature, 213, 789

Palumbo, D. C. M., Doeleman, S. S., Johnson, M. D., Bouman, K. L., & Chael, A. A. 2019, The Astrophysical Journal, 881, 62

Palumbo, D. C. M., Wong, G. N., & Prather, B. S. 2020, The Astrophysical Journal, 894, 156

Pandya, A., Chandra, M., Joshi, A., & Gammie, C. F. 2018, The Astrophysical Journal, 868, 13

Pandya, A., Zhang, Z., Chandra, M., & Gammie, C. F. 2016, The Astrophysical Journal, 822, 34

Papaloizou, J. C. B., & Pringle, J. E. 1984, Monthly Notices of the Royal Astronomical Society, 208, 721

Parfrey, K., Philippov, A., & Cerutti, B. 2019, Phys. Rev. Lett., 122, 035101

Park, J., Hada, K., Kino, M., et al. 2019, ApJ, 871, 257

Penna, R. F., McKinney, J. C., Narayan, R., et al. 2010, Monthly Notices of the Royal Astronomical Society, 408, 752

Penrose, R., & Floyd, R. M. 1971, Nature Physical Science, 229, 177

Phinney, E. S. 1983, PhD Thesis Univ Camb. 1983

Poincaré, H. 1906, Rendiconti del Circolo matematico di Palermo, Vol. 21, pp. 129-176, 21, 129

Ponti, G., Morris, M. R., Terrier, R., & Goldwurm, A. 2013, Cosmic Rays in Star-Forming Environments, 34, 331

Porth, O., Mizuno, Y., Younsi, Z., & Fromm, C. M. 2021, Monthly Notices of the Royal Astronomical Society, 502, 2023

Porth, O., Olivares, H., Mizuno, Y., et al. 2017, Computational Astrophysics and Cosmology, 4, 1

Porth, O., Chatterjee, K., Narayan, R., et al. 2019, ApJS, 243, 26

Prieto, M. A., Fernández-Ontiveros, J. A., Markoff, S., Espada, D., & González-Martín, O. 2016, Monthly Notices of the Royal Astronomical Society, 457, 3801

Pringle, J. E., & Rees, M. J. 1972, Astronomy and Astrophysics, 21, 1

Psaltis, D., & Johannsen, T. 2012, The Astrophysical Journal, 745, 1

Pu, H.-Y., Wu, K., Younsi, Z., et al. 2017, ApJ, 845, 160

Qian, Q., Fendt, C., Noble, S., & Bugli, M. 2016, ApJ, 834, 29

Quataert, E. 1998, The Astrophysical Journal, 500, 978

Quataert, E., & Gruzinov, A. 1999, The Astrophysical Journal, 520, 248

—. 2000, The Astrophysical Journal, 545, 842

Quataert, E., & Narayan, R. 1999, The Astrophysical Journal, 520, 298

Rees, M. J. 1984, Annual Review of Astronomy and Astrophysics, 22, 471

Rees, M. J., Begelman, M. C., Blandford, R. D., & Phinney, E. S. 1982, Nature, 295, 17

Reid, M. J., Schmitt, J. H. M. M., Owen, F. N., et al. 1982, The Astrophysical Journal, 263, 615

Reissner, H. 1916, Ann. Phys., 355, 106

Remillard, R. A., & McClintock, J. E. 2006, Annual Review of Astronomy and Astrophysics, 44, 49

Ressler, S. M., Quataert, E., & Stone, J. M. 2018, Monthly Notices of the Royal Astronomical Society, 478, 3544

—. 2020a, Monthly Notices of the Royal Astronomical Society, 492, 3272

Ressler, S. M., Tchekhovskoy, A., Quataert, E., Chandra, M., & Gammie, C. F. 2015, Monthly Notices of the Royal Astronomical Society, 454, 1848

Ressler, S. M., Tchekhovskoy, A., Quataert, E., & Gammie, C. F. 2017, Monthly Notices of the Royal Astronomical Society, 467, 3604

Ressler, S. M., White, C. J., Quataert, E., & Stone, J. M. 2020b, The Astrophysical Journal Letters, 896, L6

Reynolds, C. S., Di Matteo, T., Fabian, A. C., Hwang, U., & Canizares, C. R. 1996, Monthly Notices of the Royal Astronomical Society, 283, L111

Rezzolla, L., & Zanotti, O. 2013, Relativistic Hydrodyn.

Ricarte, A., Prather, B. S., Wong, G. N., et al. 2020, Monthly Notices of the Royal Astronomical Society, 498, 5468

Rieger, F. M. 2019, Galaxies Vol 7 Issue 3 P 78, 7, 78

Ripperda, B., Bacchini, F., & Philippov, A. A. 2020, Astrophys. J., 900, 100

Ripperda, B., Porth, O., Sironi, L., & Keppens, R. 2019, Monthly Notices of the Royal Astronomical Society, 485, 299

Rout, S. K., Vadawale, S., & Méndez, M. 2020, The Astrophysical Journal Letters, 888, L30

Rowan, M. E., Sironi, L., & Narayan, R. 2017, The Astrophysical Journal, 850, 29

Ruderman, M. A., & Sutherland, P. G. 1975, Astrophys. J., 196, 51

Ruffini, R., & Wheeler, J. A. 1971, Physics Today, 24, 30

Rusanov, V. 1962, USSR Computational Mathematics and Mathematical Physics, 1, 304

Russell, H. R., Fabian, A. C., McNamara, B. R., & Broderick, A. E. 2015, Monthly Notices of the Royal Astronomical Society, 451, 588

Ruszkowski, M., & Begelman, M. C. 2002, The Astrophysical Journal, 573, 485

Ryan, B. R., & Dolence, J. C. 2020, The Astrophysical Journal, 891, 118

Ryan, B. R., Dolence, J. C., & Gammie, C. F. 2015, The Astrophysical Journal, 807, 31

Ryan, B. R., Dolence, J. C., Gammie, C. F., Ressler, S. M., & Miller, J. 2019, Astrophysics Source Code Library, ascl:1909.007

Ryan, B. R., Ressler, S. M., Dolence, J. C., Gammie, C., & Quataert, E. 2018, The Astrophysical Journal, 864, 126

Ryan, B. R., Ressler, S. M., Dolence, J. C., et al. 2017, The Astrophysical Journal Letters, 844, L24

Rybicki, G. B., & Lightman, A. P. 1979, New York, Wiley-Interscience, 1979. 393 p.

Sądowski, A., & Gaspari, M. 2017, Monthly Notices of the Royal Astronomical Society, 468, 1398

Sądowski, A., Narayan, R., McKinney, J. C., & Tchekhovskoy, A. 2014, Monthly Notices of the Royal Astronomical Society, 439, 503

Sądowski, A., Narayan, R., Penna, R., & Zhu, Y. 2013a, Monthly Notices of the Royal Astronomical Society, 436, 3856

Sądowski, A., Narayan, R., Tchekhovskoy, A., & Zhu, Y. 2013b, Monthly Notices of the Royal Astronomical Society, 429, 3533

Sądowski, A., Wielgus, M., Narayan, R., et al. 2017, Monthly Notices of the Royal Astronomical Society, 466, 705

Saida, H. 2017, Progress of Theoretical and Experimental Physics, 2017, 053E02

Salpeter, E. E. 1964, The Astrophysical Journal, 140, 796

Sargent, W. L. W., Young, P. J., Boksenberg, A., et al. 1978, The Astrophysical Journal, 221, 731

Schmidt, M. 1963, Nature, 197, 1040

Schödel, R., Eckart, A., Mužić, K., et al. 2007, Astronomy and Astrophysics, 462, L1

Schödel, R., Morris, M. R., Muzic, K., et al. 2011, Astronomy and Astrophysics, 532, A83

Schödel, R., Ott, T., Genzel, R., et al. 2002, Nature, 419, 694

Schwarzschild, K. 1916, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), 1916, Seite 189-196

Sen, A. K. 1964, The Physics of Fluids, 7, 1293

Shakura, N. I., & Sunyaev, R. A. 1973, Astronomy and Astrophysics, 24, 337

Shapiro, S. L., Lightman, A. P., & Eardley, D. M. 1976, The Astrophysical Journal, 204, 187

Sharma, P., Hammett, G. W., Quataert, E., & Stone, J. M. 2006, ApJ, 637, 952

Sharma, P., Quataert, E., Hammett, G. W., & Stone, J. M. 2007, The Astrophysical Journal, 667, 714

Shcherbakov, R. V. 2008, The Astrophysical Journal, 688, 695

Shcherbakov, R. V., Penna, R. F., & McKinney, J. C. 2012, The Astrophysical Journal, 755, 133

Shepherd, M. C. 1997, Astron. Data Anal. Softw. Syst. VI, 125, 77

Shiokawa, H. 2013, PhD Thesis

Shivamoggi, B. K. 1981, Appl. Sci. Res., 37, 291

Sironi, L., Rowan, M. E., & Narayan, R. 2021, ApJL, 907, L44

Smirnov, O. M. 2011, Astronomy and Astrophysics, 527, A106

Soltan, A. 1982, Monthly Notices of the Royal Astronomical Society, 200, 115

Spitzer, L. 1962, Physics of Fully Ionized Gases, New York: Interscience (2nd edition), 1962

Stawarz, Ł., Aharonian, F., Kataoka, J., et al. 2006, Monthly Notices of the Royal Astronomical Society, 370, 981

Stepney, S., & Guilbert, P. W. 1983, Mon. Not. R. Astron. Soc., 204, 1269

Straub, O., Vincent, F. H., Abramowicz, M. A., Gourgoulhon, E., & Paumard, T. 2012, Astron. Astrophys., 543, A83

Sturrock, P. A. 1971, Astrophys. J., 164, 529

Svensson, R., & Zdziarski, A. A. 1989, Ultra-Hot Plasmas Electron-Positron Pairs Astrophys., 1

Takahashi, K., Toma, K., Kino, M., Nakamura, M., & Hada, K. 2018, ApJ, 868, 82

Takahashi, M., Nitta, S., Tatematsu, Y., & Tomimatsu, A. 1990, Astrophys. J. V363 P206, 363, 206

Takahashi, R. 2004, The Astrophysical Journal, 611, 996

Tanabe, K., & Nagataki, S. 2008, Phys. Rev. D, 78, 024004

Tchekhovskoy, A., & McKinney, J. C. 2012, Monthly Notices of the Royal Astronomical Society, 423, L55

Tchekhovskoy, A., Narayan, R., & McKinney, J. C. 2010, Astrophys. J., 711, 50

—. 2011, Monthly Notices of the Royal Astronomical Society, 418, L79

Teo, E. 2003, General Relativity and Gravitation, 35, 1909

The EHT MWL Science Working Group, Algaba, J. C., Anczarski, J., et al. 2021, ApJL, 911, L11

Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2017, Interferometry and Synthesis in Radio Astronomy, by A. Richard Thompson, James M. Moran, and George W. Swenson, Jr. 3rd ed. Springer, 2017., doi:10.1007/978-3-319-44431-4

Thompson, C. 2019, The Astrophysical Journal, 874, 48

Totani, T. 2006, Publications of the Astronomical Society of Japan, 58, 965

Tóth, G. 2000, Journal of Computational Physics, 161, 605

Tsukamoto, N., Li, Z., & Bambi, C. 2014, Journal of Cosmology and Astroparticle Physics, 06, 043

van der Gucht, J., Davelaar, J., Hendriks, L., et al. 2020, A&A, 636, A94

Vincent, F. H., Wielgus, M., Abramowicz, M. A., et al. 2021, Astronomy and Astrophysics, 646, A37

Volonteri, M., Madau, P., Quataert, E., & Rees, M. J. 2005, The Astrophysical Journal, 620, 69

Walker, R. C., Hardee, P. E., Davies, F. B., Ly, C., & Junor, W. 2018, The Astrophysical Journal, 855, 128

Walsh, J. L., Barth, A. J., Ho, L. C., & Sarzi, M. 2013, The Astrophysical Journal, 770, 86

Wang, C.-C., & Zhou, H.-Y. 2009, Monthly Notices of the Royal Astronomical Society, 395, 301

Watson, M., & Nishikawa, K.-I. 2010, Computer Physics Communications, 181, 1750

Webster, B. L., & Murdin, P. 1972, Nature, 235, 37

Werner, G. R., Uzdensky, D. A., Begelman, M. C., Cerutti, B., & Nalewajko, K. 2018, Monthly Notices of the Royal Astronomical Society, 473, 4840

White, C. J., Quataert, E., & Blaes, O. 2019, The Astrophysical Journal, 878, 51

White, C. J., Stone, J. M., & Gammie, C. F. 2016, Astrophys. J. Suppl. Ser., 225, 22

Wielgus, M., Horák, J., Vincent, F., & Abramowicz, M. 2020, Physical Review D, 102, 084044

Witzel, G., Eckart, A., Bremer, M., et al. 2012, The Astrophysical Journal Supplement Series, 203, 18

Wollman, E. R., Geballe, T. R., Lacy, J. H., Townes, C. H., & Rank, D. M. 1977, The Astrophysical Journal Letters, 218, L103

Woltjer, L. 1966, The Astrophysical Journal, 146, 597

Wong, G. N. 2021, ApJ, 909, 217

Wong, G. N., Du, Y., Prather, B. S., & Gammie, C. F. 2021a, ApJ, 914, 55

Wong, G. N., Ryan, B. R., & Gammie, C. F. 2021b, The Astrophysical Journal, 907, 73

Wu, M.-C., Xie, F.-G., Yuan, Y.-F., & Gan, Z. 2016, Monthly Notices of the Royal Astronomical Society, 459, 1543

Yang, H., Nichols, D. A., Zhang, F., et al. 2012, Physical Review D, 86, 104006

Yarza, R., Wong, G. N., Ryan, B. R., & Gammie, C. F. 2020, The Astrophysical Journal, 898, 50

Yoon, D., Chatterjee, K., Markoff, S. B., et al. 2020, Monthly Notices of the Royal Astronomical Society, 499, 3178

Younsi, Z., Porth, O., Mizuno, Y., Fromm, C. M., & Olivares, H. 2020, Proc. Int. Astron. Union, 14, 9

Younsi, Z., Zhidenko, A., Rezzolla, L., Konoplya, R., & Mizuno, Y. 2016, Physical Review D, 94, 084025

Yu, Z., Yuan, F., & Ho, L. C. 2010, ApJ, 726, 87

Yuan, F., Gan, Z., Narayan, R., et al. 2015, ApJ, 804, 101

Yuan, F., Markoff, S., Falcke, H., & Biermann, P. L. 2002, Astronomy and Astrophysics, 391, 139

Yuan, F., & Narayan, R. 2014, Annual Review of Astronomy and Astrophysics, 52, 529

Yuan, F., Quataert, E., & Narayan, R. 2003, The Astrophysical Journal, 598, 301

Zdziarski, A. A. 1985, Astrophys. J. Part 1 ISSN 0004-637X Vol 289 Feb 15 1985 P 514-525, 289, 514

Zhang, W., Dovčiak, M., & Bursa, M. 2019, The Astrophysical Journal, 875, 148

Zhdankin, V., Uzdensky, D. A., Werner, G. R., & Begelman, M. C. 2019, Physical Review Letters, 122, 055101