A Universal Power-law Prescription for Variability from Synthetic Images of Black Hole Accretion Flows

Georgiev, B.; Chatterjee, K.; Yoon, D.; Porth, O.; Markoff, S.; Musoke, G.; Event Horizon Telescope Collaboration

DOI
10.3847/2041-8213/ac65eb

Publication date
2022

Document Version
Final published version

Published in
Astrophysical Journal Letters

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
A Universal Power-law Prescription for Variability from Synthetic Images of Black Hole Accretion Flows

Abstract

We present a framework for characterizing the spatiotemporal power spectrum of the variability expected from the horizon-scale emission structure around supermassive black holes, and we apply this framework to a library of general relativistic magnetohydrodynamic (GRMHD) simulations and associated general relativistic ray-traced images relevant for Event Horizon Telescope (EHT) observations of Sgr A*. We find that the variability power spectrum is generically a red-noise process in both the temporal and spatial dimensions, with the peak in power occurring on the longest timescales and largest spatial scales. When both the time-averaged source structure and the spatially integrated light-curve variability are removed, the residual power spectrum exhibits a universal broken power-law behavior. On small spatial frequencies, the residual power spectrum rises as the square of the spatial frequency and is proportional to the variance in the centroid of emission. Beyond some peak in variability power, the residual power spectrum falls as that of the time-averaged source structure, which is similar across simulations; this behavior can be naturally explained if the variability arises from a multiplicative random field that has a steeper high-frequency power-law index than that of the time-averaged source structure. We briefly explore the ability of power spectral variability studies to constrain physical parameters relevant for the GRMHD simulations, which can be scaled to provide predictions for black holes in a range of systems in the optically thin regime. We present specific expectations for the behavior of the M87+ and Sgr A+ accretion flows as observed by the EHT.

Unified Astronomy Thesaurus concepts: Accretion (1578); Magnetohydrodynamics (1964); Galactic center (565); Black holes (162); Very long baseline interferometry (1769)

1. Introduction

The Event Horizon Telescope (EHT) has opened a new window on supermassive black holes. With the advent of millimeter and submillimeter very long baseline interferometry (VLBI), it has become possible to image a handful of targets with resolutions sufficient to resolve the lensed event horizon, silhouetted against the luminous plasma in its vicinity. Images of the two horizon-scale EHT targets, M87+ and Sgr A+, have now been published, revealing in unprecedented detail the astrophysical processes at work in the innermost regions of accretion flows and in the region responsible for launching jets (Event Horizon Telescope Collaboration et al. 2019a, 2019b, 2019c, 2019d, 2019e, 2019f, hereafter M87+ Paper I–VI; Event Horizon Telescope Collaboration et al. 2022a, 2022b, 2022c, 2022d, 2022e, 2022f, hereafter Papers I–VI).

Images of both M87+ and Sgr A+ are grossly consistent with the expectations from theoretical modeling of the near-horizon region (Luminet 1979; Falcke et al. 2000; Broderick & Loeb 2006; Broderick et al. 2009; Broderick & Loeb 2009; Mościbrodzka et al. 2009; Dexter et al. 2009, 2012; Mościbrodzka et al. 2014, 2016; M87+ Paper V; Paper V). This is largely a result of two generic features that shape horizon-scale images: the strong gravitational lensing that imposes the central flux depression, or shadow, independent of the underlying accretion flow; and the relativistic motions within the emission region due to Doppler beaming and shifting (Luminet 1979; Falcke et al. 2000; Broderick & Loeb 2009; Narayan et al. 2019). The former is fully described by general relativity and thus fully determined in the absence of new physics. In contrast, the latter is dependent on astrophysical details within the emission region and thus model dependent.

It is, therefore, nontrivial that state-of-the-art general relativistic magnetohydrodynamic (GRMHD) simulations of the magnetized plasmas near the black hole are in broad agreement with the observed image morphologies (M87+ Paper V; Paper V). For example, in the case of M87+, the dominance of the emission in the south despite the jet extending primarily westward is a prediction of spin-driven jet-launching models, in which the emission region is relativistically rotating (Broderick & Loeb 2009; M87+ Paper V).

Future instruments will substantially expand the spatial and temporal scales that such images probe. Upcoming ground-based projects, e.g., the ngEHT, promise to produce improvements in resolution by pushing toward shorter wavelengths, major improvements in sensitivity, and the capacity to probe variability in these sources on timescales of minutes (see, e.g., Doeleman et al. 2019; Raymond et al. 2021). Space-based VLBI missions, building on the demonstrated recent success of RadioAstron (Gómez et al. 2016; Kardashev et al. 2017; Johnson et al. 2021), have the potential to increase imaging resolution by orders of magnitude, both improving the view of EHT targets and expanding the number of horizon-scale sources commensurately (Johnson et al. 2019; Pesce et al. 2021). Thus, it is now relevant to develop the theoretical expectations for the short-timescale (\(\lesssim GM/c^3\)) and small spatial scale (\(\lesssim GM/c^2\)) aspects of these images, where \(M\) is the mass of the supermassive black hole.\(^{151}\)

Black hole accretion flows are necessarily dynamic and nonlaminar; the MHD turbulence that facilitates the transfer of angular momentum responsible for driving material inward is an essential feature of accretion flow models on compact objects (Shakura & Sunyaev 1973; Balbus & Hawley 1991).

\(^{151}\) Hereafter, \(G\) and \(c\) are Newton’s gravitational constant and the speed of light, respectively.
The resulting variability is expected to manifest on a wide variety of temporal and spatial scales, ranging from the cyclotron scale to the orbital scale. An a priori understanding of the properties of fully developed MHD turbulence remains an active area of study (see, e.g., White et al. 2020; Narayan et al. 2022, and references therein). Misaligned accretion flows, i.e., systems in which the angular momentum of the accreting gas is not aligned with spin of the central black hole, will result in additional variability (Dexter & Fragile 2013; Chatterjee et al. 2020; Ressler et al. 2020).

This variability is already observed in unresolved observations of accreting black holes across the mass spectrum, ranging from X-ray binaries to active galactic nuclei (see, e.g., Wagner & Witzel 1995; Remillard & McClintock 2006). For Sgr A* in particular, the light curve is significantly variable and has been well characterized at near-infrared and radio wavelengths (Wielgus et al. 2022, and references therein; Witzel et al. 2012; Gravity Collaboration et al. 2020; Witzel et al. 2021; Paper II). Below a break that lies between 2 and 8 hr, the light-curve variability is consistent with a red-noise process, having a power-law spectrum that is dominated by the longest timescales. At 1.3 mm wavelengths, the typical degree of variability is observed to be less than 10% during the 2017 EHT observing campaign (Wielgus et al. 2022). The EHT observations of Sgr A* imply a similar degree of variability in the source structure (Paper IV).

The variability properties of Sgr A* are grossly consistent with the expectations from GRMHD simulations (Paper V). These simulations exhibit coherent and stochastic structural fluctuations on timescales spanning all of those accessible to the simulation domain; for Sgr A*, this ranges from seconds to many days.

The presence of substantial horizon-scale variability in EHT targets has dramatic implications for the analysis of EHT VLBI data. Sgr A* and, to a lesser degree, M87+ necessarily violate the stationarity assumptions underlying Earth-aperture synthesis. In the case of the former, this presents a significant impediment to imaging and model comparison (Broderick et al. 2022; Paper III; Paper IV).

At the same time, direct measurements of the structural variability—i.e., variability caused by the evolving source structure—provide a novel constraint on models of the underlying astrophysical processes responsible for accretion (Wielgus et al. 2020).

Here we explore the properties of the structural variability observed within simulated images from the large library of GRMHD simulations produced to interpret Sgr A* (Paper V). These simulations span a wide range of black hole and accretion flow properties, including spin, inclination, flow type, and microphysical electron heating prescription. We do this primarily via the construction of statistical measures of the structural variability within the images. We find a surprising degree of uniformity, despite the extreme variation among image morphologies.

We begin in Section 2 with a description of the GRMHD simulations. Definitions of the statistical measures employed can be found in Section 3. Application to the GRMHD library from Paper V can be found in Section 4, and the dependence on the accretion flow properties is discussed in Section 5. Finally, conclusions may be found in Section 6.

2. Horizon-scale Observations and Simulations of Sgr A*

The immediate application of this work is to observations of the supermassive black hole at the Galactic center, Sgr A*. In this section, we describe the current leading-edge observations of Sgr A* using the EHT and identify the necessary outputs any physical model must contain. Then, we compile and describe a library of GRMHD simulations, from which the synthetic EHT observables are created.

2.1. EHT Observables

The EHT is a global VLBI array of radio telescopes. Each pair of concurrently observing telescopes forms a baseline and measures the complex Fourier transform of the on-sky intensity distribution $I(x, y, t)$—visibility $\hat{V}(u, v, t)$. Hence, at time $t$, the measured visibility is

$$\hat{V}(u, v, t) = \int dx dy e^{-2\pi i (u x + v y)} I(x, y, t).$$  (1)

In 2017, the EHT array consisted of eight stations at six geographical sites (Paper II). However, not all stations can observe Sgr A* simultaneously, and hence the available data correspond to only 13 nonredundant baselines. This results in a sparse sampling in the Fourier domain. The Fourier coordinates $(u, v)$ probed by each baseline are time dependent themselves, as an effect of the Earth rotation. In a standard approach, the source is assumed to be static on the timescale of the observations (~several hours), thus allowing the use of Earth’s rotation for the aperture synthesis (Thompson et al. 2017). Sgr A* constitutes an unusual case in which this condition is violated, as the angular resolution of the EHT (about 25 $\mu$as; Paper III) corresponds to a linear scale of only 2.5 Schwarzschild radii, given Sgr A*’s mass and distance (Do et al. 2019; Gravity Collaboration et al. 2019). This translates to a light-crossing time of 100 s, or $5 GM/c^2$, and sets the shortest timescale of structural variability to which EHT observations are sensitive. The longest timescales probed during an observing night correspond to several hours, or ~$10^5 GM/c^2$. Therefore, the source is expected to vary over the timescale of the observations.

The 2017 EHT VLBI baselines probe spatial scales varying from 0.5 to 8.7 GA, corresponding to angular scales of about 25–400 $\mu$as (Paper III). Additionally, two of the EHT sites, the Atacama Large Millimeter/submillimeter Array (ALMA) and the Submillimeter Array (SMA), are themselves connected element interferometric arrays, probing linear scales of ~$10^6$ Schwarzschild radii. While the phased-up signal is used for the VLBI EHT observations (Goddi et al. 2019; Paper II), the internal correlations between the individual telescopes can be used to generate the total flux light curves, representing the time-dependent total flux of the compact Sgr A* source (Wielgus et al. 2022), unresolved at spatial scales of ~100 k\lambda. These observations effectively correspond to $V(0, 0, t)$ for a compact Sgr A* source. We neglect any large-scale structure like the Sgr A* “minispiral,” which is overresolved on baselines longer than ~100 k\lambda and filtered out from the observations (any large-scale source structure is naturally missing in the simulated images, which only represent the event horizon scale compact source and are limited by a correspondingly narrow field of view (FOV). The variability of the light curve thus corresponds to scaling of the integrated flux.
density of the source, rather than to a structural variability related to spatial redistribution of brightness in the resolved image.

In this work we characterize the spatiotemporal variability of the visibility domain representation of a broad collection of GRMHD simulations. We attempt to characterize the typical deviation of $\mathcal{V}(u, v, t)$ from the average across the whole simulation, to provide a physically motivated prior for the measured source variability probed by the EHT (Broderick et al. 2022). In order to facilitate the structural variability analysis of the resolved source, we consider the properties of simulated images normalized by the corresponding light curves (instantaneous total flux values).

### 2.2. Modeling the Galactic Center

In order to generate synthetic images for the black hole systems targeted by the EHT, we perform fluid simulations of accreting plasma in the Kerr black hole spacetime and simulate emission and absorption to generate the images that would be seen by a distant observer. The fluid simulations treat the plasma as an ideal, single-temperature fluid that is governed by the equations of ideal GRMHD. A detailed description of these equations can be found in Gammie et al. (2003). Our ideal GRMHD simulations make several assumptions. First, we assume that the plasma can be treated as a fluid, even though the mean free paths of particles in the flows are often much longer than the characteristic length scales in the systems. Second, we assume that the plasma has infinite conductivity, such that the magnetic field lines are "frozen" into the fluid.

In reality, kinetic processes in the (near-)collisionless plasma may increase the effective particle collision rate and allow deviations from the ideal fluid picture that can modulate typical variability timescales (see, e.g., Foucart et al. 2016). Deviations from the infinite conductivity ideal fluid approximation (including the effects of viscous or resistive dissipation and heat conduction) may alter the thermodynamics of the flow (see Foucart et al. 2017, for a discussion; see Chandra et al. 2017; Most & Noronha 2021; Most et al. 2021, for descriptions of viscosity in GRMHD; see, e.g., Ripperda et al. 2020, for an exploration of the effect of finite resistivity, especially in the context of plasmoid-mediated reconnection powering flares). It is particularly important to resolve the spatial scales on which the aforementioned dissipation mechanisms act (e.g., Ripperda et al. 2022). Variability timescales governed by turbulent dissipation or reconnection may, however, be very different in collisionless plasmas and be strongly affected by nonthermal electron distribution functions, requiring a first-principles relativistic kinetic approach (e.g., Guo et al. 2014; Sironi & Spitkovsky 2014; Kunz et al. 2014; Werner et al. 2016; Zhdamkin et al. 2017; Comiso & Sironi 2018; Bransgrove et al. 2021; Nättäli & Beloborodov 2021). All of these effects could cause significant changes in variability not captured in the simulations used in this work.

We also assume that the radiation is not dynamically important to the flow, so that it can be treated entirely in a post-processing step. The nonradiative assumption is a valid approximation for our system of interest, Sgr A*, where accretion is radiatively inefficient and $M \ll M_{\text{Edd}}$; here $M$ is the accretion rate and $M_{\text{Edd}}$ is the Eddington accretion rate (Yuan & Narayan 2014). We assume an ideal gas equation of state for the plasma with adiabatic index, $\Gamma_{\text{ad}}$. The system evolves on a stationary background spacetime that is unaffected by the accretion flow, whose total energy content is orders of magnitude lower than the central supermassive black hole. The Kerr metric has two parameters: the mass of the black hole, $M$, and its dimensionless spin, $a_*=Jc/GM^2$, which in general has both an amplitude and an orientation. In this work, we choose the spin axis to point in the $y$-direction in the image and $v$-direction in the Fourier domain.

Concretely, the numerical simulations are initialized with an axisymmetric hydrodynamically stable Fishbone–Moncrief torus (Fishbone & Moncrief 1976), which is parameterized by the inner radius of the disk, $r_{\text{in}}$, and the radius at pressure maximum, $r_{\text{max}}$. The torus is chosen to rotate in either the same ($a_*>0$) or the opposite ($a_*<0$) direction as the black hole. An axisymmetric electromagnetic potential is used as the initial condition for the magnetic field (see Wong et al. 2022, for more details), and the thermal energy of the fluid is perturbed to seed development of instabilities such as the magnetorotational instability (MRI; Balbus & Hawley 1991) and kickstart accretion. After $\gtrsim 5000 \, GM/c^3$, the inner region reaches a quasi-steady state so that the initial condition of the simulation has negligible effect on the final state and the simulation can reproduce the relatively steady-state light curve observed over the 2017 EHT observations, which exhibits variability on the scales of 10% (Paper II; Paper V). There is evidence that Sgr A* may have recently been in a period of increased accretion (see, e.g., Montero-Castaño et al. 2009), but such more complex feeding scenarios are beyond the scope of this work. Far from the black hole, the simulation is dominated by its initial condition, which supplies material to the inner turbulent flow. Since the torus is initialized with finite mass, in long-duration studies, it is possible that an appreciable fraction of the disk mass may accrete onto the black hole. This net decline may introduce an "artificial" systematic downward trend in long time series generated from the simulation. For a study on longer-timescale simulations, see Appendix A.3, and for a brief study of simulation resolution, see Appendix A.2.

The strength and structure of the magnetic field during the late-time, quasi-equilibrium phase dictate the magnetic flux, $\Phi_{\text{mag}}$, threading the event horizon. Each simulation we consider equilibrates into one of two qualitatively different modes of accretion, each of which has a characteristic value of the nondimensionalized magnetic flux, $\phi_{\text{mag}} = \Phi_{\text{mag}} / \sqrt{M}$. When $\phi_{\text{mag}} \sim \phi_*$, where $\phi_* \sim 15$ (Tchekhovskoy et al. 2011, but using the unit convention of Porth et al. 2019) is the critical value at which the local magnetic pressure near the event horizon balances the inward fluid ram pressure, we obtain a magnetically arrested disk (MAD; Bisnovatyi-Kogan & Ruzmaikin 1974; Igumenshchev et al. 2003; Narayan et al. 2003). MAD accretion is choppy and characterized by transient ejection events where bubbles of magnetic flux that have accumulated on the horizon are expelled by the system. The strong magnetic pressures near the event horizon in MAD systems tend to disallow steady disk accretion, forcing the infalling plasma to accrete in thin, disordered accretion filaments or strands. In contrast, when $\phi_{\text{mag}} \ll \phi_*$, the system produces a standard and normal evolution (SANE) disk (Narayan et al. 2012; Sadowski et al. 2013). SANE disks are turbulent but relatively well ordered, and angular momentum transport is thought to be governed by the MRI. Because MAD and SANE accretion flows have such different morphological features, the geometry of their hot emission regions may vary drastically. A full list of fluid simulations used in this work can be found in Table 1.
The GRMHD fluid simulations are used to generate images through general relativistic radiative transfer (GRRT), which proceeds in two steps. First, photon trajectories are computed backward from the camera into the simulation domain by solving the geodesic equation. Then, the covariant radiative transfer equation is solved forward to the camera along the precomputed photon geodesics, accounting for emission and absorption from the local plasma. The majority of results in this work consider images generated under the “fastlight” assumption, in which we approximate the fluid as unchanging during the entire GRRT process so that any image corresponds to a single (Kerr–Schild) time slice. We briefly discuss differences between fastlight and the “slowlight” alternative later (Section A.1).

The GRRT procedure takes as input the following:

1. the fluid and electromagnetic data (mass density, fluid internal energy, fluid velocity, magnetic field) on a particular Kerr–Schild time slice;
2. a prescription for translating local fluid data into emission and absorption transfer coefficients at the appropriate observing frequency;
3. the mass and spin of the central black hole; and
4. camera parameters (including distance to source D and inclination angle i).

Here the inclination is defined as the angle between the line of sight to the observer and the black hole spin axis. In principle, we could introduce an azimuthal position angle of the camera, but there is not a unique choice for accretion flows aligned with the black hole spin.

The equations of nonradiative GRMHD are invariant under rescalings of length and time with the black hole mass, $M$, and rescalings of the fluid density, internal energy, and square of magnetic field strength with some density scale. The angular scale of the image, $GM/c^2$, is fixed by providing the distance to the source, $D$. Following Paper V, we set the mass of the black hole to be $M = 4.14 \times 10^6 M_\odot$, where $M_\odot$ is the solar mass, and the distance to be $D = 8.127$ kpc (Do et al. 2019; Gravity Collaboration et al. 2019). The FOV of the images is chosen large enough to contain the majority of the flux (Paper V). The density scale is chosen so that the average simulated observed compact flux at 230 GHz matches the EHT-observed value of 2.4 Jy (Wielgus et al. 2022).\footnote{Due to approximations made in fitting for the density scale, the total flux in the simulations varies by about 20%.}

For small enough mass accretion rates, the gas is optically thin and the image intensity is proportional to the density scale. At 2.4 Jy, Sgr A’ is marginally optically thin, so these GRMHD images can loosely approximate other optically thin black hole systems, such as M87\footnote{Golden}, with an appropriate mass and flux rescaling.

The expected near-collisionless nature of the plasma in the Galactic center accretion system likely produces a flow in which the electron and ion components are best described as having different temperatures, $T_e \approx T_i$. Since our fluid simulations only track the total energy of the fluid, it is necessary to prescribe the temperature of the electrons, which are responsible for the emission. We use the electron temperature prescription described in Paper V, which is based on the model of Mościbrodzka et al. (2016), in which the ratio of ion to electron temperatures is a function of the local plasma β = $P_{\text{gas}}/P_{\text{mag}}$ according to

$$\frac{T_i}{T_e} = \frac{R_{\text{high}}}{1 + \beta^2} + \frac{1}{1 + \beta^2}. \quad (2)$$

Heuristically, this model allows the two species to have approximately equal temperatures in the low-density, highly magnetized jet region and produces increasingly cool electrons in the bulk disk as the free parameter $R_{\text{high}}$ is increased. In this work, we assume that the electrons are well described by a thermal distribution, which is parameterized by a single (temperature) variable.

A more comprehensive discussion of the steps and approximations involved in the generation of synthetic images can be found in Paper V and Wong et al. (2022). In the present work we consider fluid simulations performed by the following GRMHD codes: KHarMa and iHarM3D (Prather et al. 2021), KoRAL (Sadowski et al. 2013, 2014), and BHAC (Porth et al. 2017). The synthetic images of the GRMHD “snapshots” are generated by performing general relativistic ray-tracing.

These resulting images or “frames” are the input for the analysis in this work. We consider images produced using the ipole and BHOSs GRRT codes; Gold et al. (2020) present a comparison of contemporary GRRT codes. See Porth et al. (2019) for a comparison of contemporary GRMHD codes.

### 2.2.1. HARM

iHarM3D is an implementation of the HARM algorithm outlined in Gammie et al. (2003), and KHarMa is a GPU port of iHarM3D closely following the original codebase and

<table>
<thead>
<tr>
<th>Table 1: GRMHD Fluid Simulation Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Label</strong></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
</tbody>
</table>

Note. $^a$ In units of the gravitational radius, $GM/c^2$.  

R. D. BHAC KORAL

KoRAL and BHAC are generated by performing general relativistic ray-tracing. These resulting images or “frames” are the input for the analysis in this work. We consider images produced using the ipole and BHOSs GRRT codes; Gold et al. (2020) present a comparison of contemporary GRRT codes. See Porth et al. (2019) for a comparison of contemporary GRMHD codes.
functionality (see also Prather et al. 2021). The GRMHD models produced by KHARMA that are considered for this work include MAD and SANE states, each spanning five spins, $a_*=0, \pm 1/2, \pm 15/16$ (hereafter written as 0, ±0.5, and ±0.94). These simulations form model set “A.” A comprehensive discussion of these models can be found in V. Dhruv (2022, in preparation).

The GRMHD snapshots produced by these codes are imaged using the publicly available GRRT code ipole (Mościbrodzka & Gammie 2018). A comprehensive discussion of the steps and approximations involved in the generation of these synthetic images from HARM output using ipole can be found in Wong et al. (2022). The images span nine inclinations from 10° to 170° and four values of $R_{\text{high}}$ from 1 to 160. From each simulation, three windows, each of length 5000 $\text{GM}/c^3$, are chosen and treated as separate simulations with different density scales.

We also consider a high-cadence MAD simulation ($a_*=15/16$) carried out with iharm3d, for the study of short timescales and the effects of the fastlight approximation (see model set “B”).

### 2.2.2. KORAL

The GRMHD models used from KORAL are all in the MAD state with spins of $a_*=0, \pm 0.3, \pm 0.5, \pm 0.7, \text{and } \pm 0.9$. These simulations were run for 100,000 $\text{GM}/c^3$ to study the effects of long-timescale trends. They form model set “C” and have been studied in Narayan et al. (2022).

These simulations were also imaged using ipole (Mościbrodzka & Gammie 2018), using a cadence of 100 $\text{GM}/c^3$ between times 10,000 $\text{GM}/c^3$ and $\sim$100,000 $\text{GM}/c^3$, with some segments of some simulations further imaged at lower cadence. Using the electron temperature prescription of Mościbrodzka et al. (2016), only $R_{\text{high}}=20$ is sampled. We sample inclinations of $i \in \{10, 30, 50, 70, 90, 110, 130, 150, 170\}$. For each model set, we set the density scale, $\mathcal{M}$, to produce an average total flux of $2.4 \text{ Jy}$ across the time window, consistent with the Sgr A* total flux measured by Wielgus et al. (2022) during the EHT 2017 campaign. Since these simulations are run for such a long time, the mass accretion rate drops noticeably as the torus drains. To counteract this, we allow the density scaling $\mathcal{M}$ to evolve slowly with time by fitting for $a$ and $b$ in $\ln \mathcal{M} = a + b(t c^3/\text{GM})$ instead of the constant value of $\mathcal{M}$ used for all other codes in this work. This procedure effectively assumes that the effects of the draining torus are limited only to the total flux light curve; this assumption is somewhat tested in Section A.3.

#### 2.2.3. BHAC

The Black Hole Accretion Code (BHAC) is a multidimensional GRMHD code to solve the equations of GRMHD in arbitrary spacetimes (Porth et al. 2017; Olivares et al. 2019). GRMHD simulations of magnetized accretion flows onto a black hole have been performed in both MAD and SANE states with five different black hole spins, $a_*=0, \pm 1/2, \pm 15/16$. Imaging of these GRMHD simulations is performed by the GRRT code (BHÖSS; Younsi et al. 2012, 2020). Only inclinations less than $90^\circ$ are considered, as the simulations are approximately equatorial symmetric. These simulations and images correspond to model set “D.”

We have also performed a resolution test by performing the same fluid simulation with three increasingly lower resolutions. This simulation is in the MAD state with $a_*=15/16$ and forms model set “E.”

### 2.3. Salient Features of GRMHD Simulations

Figure 1 shows an example output from the simulating procedure for two simulations of model set A on each row. The
left three panels show three consecutive frames, and the right panel shows an average over all frames in this simulation. The differences in the two simulations chosen are indicative of the wide range of morphology and variability present among simulations.

Source emission is limited to within several times the gravitational length scale. Images typically show asymmetric ring-like emission, with an asymmetric flux distribution extending to larger scales. Variability in this extended region is mainly composed of coherent moving features. On the short timescales shown, these features move less, i.e., the variability on these timescales is lower.

3. Power Spectral Density Characterization of Variability

For each simulation in the GRMHD library, the evolving emission structure is described by a series of “frames,” each of which is uniquely determined at time \( t \) by its intensity at every pixel location \((x, y)\). The full time series of frames from a single simulation then describes a dynamic “image” \( I(x, y, t) \) of the evolving emission structure. Our goal is to characterize the variability in the images contained within the GRMHD library. In this paper we work primarily with both spatial and temporal power spectral density (PSD) representations of the source structure. In this section, we define our PSD construction formalism and describe some general behavior. As we shall see in Section 4, we will be dealing with red-noise processes, so we also introduce filtering and averaging procedures to produce accurate statistical representations in the temporal and spatial scales of interest. Terminology and notation is summarized in Table 3.

3.1. Variability PSD Definitions

The spatiotemporal PSD of a set of images is given by the complex square of the three-dimensional Fourier transform of \( I(x, y, t) \),

\[
\Phi(u, v, \omega) = |\mathcal{F}_{xyt}[I(x, y, t)]|^2 = \mathcal{F}^*_{xyt}[I(x, y, t)]\mathcal{F}_{xyt}[I(x, y, t)],
\]

where

\[
\mathcal{F}_{xyt}(f) \equiv \iiint f(x, y, t)e^{-2\pi i(ax+by+ct)}dxdydt
\]

is a three-dimensional spatial Fourier transform from \((x, y, t)\) to \((u, v, \omega)\), and the asterisk denotes complex conjugation. We will often work with the “spatial PSD” of the image, in which the Fourier transform is only performed on the two-dimensional spatial portion of \( I(x, y, t) \),

\[
\mathcal{F}_{xy}(f) \equiv \oint f(x, y, t)e^{-2\pi i(ax+by)}dxdy.
\]

In this paper, we will often refer to the spatial PSD as simply the PSD, and we will specify otherwise when referring to the full spatiotemporal PSD.

In the absence of variability, the PSD computed at any time \( t \) would be identical to the PSD computed at any other time. However, variability in the source introduces variability into the PSD as well, which we can characterize in terms of its variance with respect to an average,

\[
\langle P_T(u, v, t) \rangle = \langle |\mathcal{F}_{xy}[I(x, y, t) - \bar{I}_T(x, y, t)]|^2 \rangle,
\]

where the angle bracket notation,

\[
\langle f(t) \rangle = \frac{1}{\tau} \int_0^\tau f(t)dt,
\]

represents an average in time over the total duration \( \tau \) of the simulation. In Equation (6), \( \bar{I}_T(x, y, t) \) is a time-averaged version of the image, defined using a Gaussian smoothing kernel,

\[
\bar{I}_T(x, y, t) = \int I(x, y, t') \frac{1}{\sqrt{2\pi T}} \exp\left[-\frac{(t - t')^2}{2T^2}\right] dt'.
\]

Here \( T \) is the smoothing timescale, and we introduce the bar and subscript notation \( \bar{I}_T \) to denote Gaussian smoothing over this timescale. As \( T \to \infty \), \( \langle P_T \rangle \) measures the variance of the amplitude spectrum at every spatial frequency \((u, v)\). For finite \( T \), contributions to this variance from timescales longer than \( T \) are suppressed. For the red-noise processes relevant for this work, \( \langle P_T \rangle \) will thus be dominated by the variability on timescale \( T \) (for more discussion and application, see Sections 3.3 and 4.2).

We note that \( \langle P_f \rangle \) is in general not a complete description of the source variability. Because physical variability originates from continuous processes, we expect the Fourier intensities at \((u, v, t)\) and at \((u + \Delta u, v + \Delta v, t + \Delta t)\) to be strongly correlated for small values of \( \Delta u, \Delta v, \) and \( \Delta t \). We also expect the strength of this correlation to decrease as these separations increase. \( \langle P_T \rangle \) is just this correlation evaluated at \( \Delta u = \Delta v = \Delta t = 0 \).

In the context of astrophysical observations, variations on the largest spatial scales \((u, v) = (0, 0)\) are captured by the light curve,

\[
L(t) = \iint I(x, y, t)dxdy,
\]

with a similar definition of \( L_T \) containing the total flux of \( I_T \). When \( L(t) \) is known, it can be used to remove the variability on the largest spatial scale by “light-curve-normalizing” the image,

\[
\bar{I}(x, y, t) \equiv \frac{I(x, y, t)}{L(t)}.
\]

The corresponding light-curve-normalized version of \( \langle P_T \rangle \) is given by

\[
\langle \hat{P}_T(u, v, t) \rangle = \langle |\mathcal{F}_{xy}[\bar{I}(x, y, t) - \bar{I}_T(x, y, t)]|^2 \rangle,
\]

with \( \bar{I}_T(x, y, t) \) defined by replacing \( I \to \bar{I} \) in Equation (8). We refer to Equation (11) as the “residual” PSD, with an analogous image-domain counterpart \( \bar{I}(x, y, t) - \bar{I}_T(x, y, t) \) that we refer to as the residual image. In practice, removing the large-scale variability in this way also removes some fraction of the variability contained on all spatial scales that are correlated with \((u, v) = (0, 0)\). The impact of this light-curve normalization procedure is explored in Section 4.4.
Practically, the angle brackets in Equation (7) are an average over all of the frames in a simulation, that is,
\[
\langle f(t) \rangle \approx \frac{\sum_i f(t_i) \Delta t}{\sum_i \Delta t},
\]
where \(\Delta t\) is the frame spacing.153 The integral in the definition of the time-smoothed image in Equation (8) is similarly replaced with a sum,
\[
I_{T}(x, y, t) \approx \frac{\sum_i I(x, y, t_i) \exp[-(t_i - t)^2/(2\tau^2)] \Delta t}{\sum_i \exp[-t_i^2/(2\tau^2)] \Delta t},
\]
where \(t_i\) is the frame time. Therefore, outside of the time region of the simulation, the Gaussian weighting kernel is set to zero, and when \(T\) approaches the length of the simulation, the time-smoothed image approaches the uniformly weighted mean.

To compute the Fourier transforms in Equations (4) and (5), we use the fast Fourier transform algorithm implemented in NumPy (Harris et al. 2020). When we later use \(\Phi\) for a qualitative comparison, we average the temporal Fourier transform over multiple time windows of a simulation to reduce the statistical error. The spatial Fourier transforms in the calculation of \(\hat{P}_T\) and \(\hat{P}_T\) of a compact source are done on images padded with a large zero array. No significant change occurs when increasing the pixel size or reducing the FOV of the original images. The time-averaged power spectra constructed as described in Equations (6) and (11) are averaged over many images and therefore many independent realizations of the underlying random processes of interest, reducing the statistical uncertainty in the variability PSDs.

3.2. PSD Behavior on Short Baselines

While the detailed structure of the PSD depends on the specifics of the source variability, we can understand its broader structural characteristics in light of a few relatively simple considerations. The PSD structure on the largest spatial scales can be described by an expansion of the PSD about \((u, v) = (0, 0)\). For any function \(f(x, y)\), its spatial Fourier transform expanded around this point is given by
\[
\mathcal{F}_x(f) \approx \int [1 - 2\pi i(u x + v y) + ...]f(x, y) \, dx \, dy.
\]
Because \(\langle P_T \rangle\) is an even function in \((u, v)\), it must flatten out on large enough spatial scales and will thus be dominated by the first term in the Fourier expansion
\[
\langle P_T \rangle \approx \langle P_T \rangle(0, 0) = \langle |L(t) - L_T(t)|^2 \rangle
\]
for small enough \((u, v)\), typically those measuring structures larger than the source size. As \(T \to \infty\), the value of \(\langle P_T \rangle\) on the largest scales approaches the variance in the light curve. For \(T \to 0\), the zero-baseline behavior is determined by the temporal PSD of the light curve; see Sections 3.3 and 4.3. However, the first term in the expansion of \(\langle P_T \rangle\) is zero by construction. Therefore, the leading term in the expansion is
\[
\langle P_T \rangle \approx 4\pi^2 \langle \int (u x + v y)(\mathcal{L} - \mathcal{L}_T) \, dx \, dy \rangle^2
\]
\[
= 4\pi^2 \langle u(\mathcal{X} - \mathcal{X}_T) + v(\mathcal{Y} - \mathcal{Y}_T) \rangle^2,
\]
and a similar definition for \(\mathcal{Y}(t)\). It follows that on the largest spatial scales \(\langle P_T \rangle\) along any direction in \((u, v)\) will behave like a power law in baseline length with an index of 2. It is also interesting to note that for a symmetric (i.e., even) function, all the odd terms of the expansion in Equation (14) are zero. Therefore, a large-scale power-law index of 2 is an indicator of asymmetric variability in the residual image. More generally, the coefficient in Equation (16) is a measure of the amount of variability in the centroid of emission.

The identification of the short-baseline variability described here with an astrometric quantity (the position of the image centroid) raises a practical question. Where an absolute phase reference is available, such a measurement is well defined (see, e.g., Reid & Honma 2014; Broderick et al. 2011). However, for the current EHT specifically and VLBI studies typically, atmospheric phase delays preclude a full reconstruction of the visibility phases on timescales longer than the atmospheric coherence time; at 230 GHz for the EHT, this is on the timescale of seconds (M87* Paper II). These unknown station-based phase errors are degenerate with an arbitrary shift in the image centroid. We discuss the implications for the short-baseline structure of \(\langle P_T \rangle\) further in Appendix B. Because our focus in this paper is in the theoretical predictions of variability and the signatures of astrophysical processes contained therein, in the remainder of this paper we presume that absolute phase information is available.

3.3. Temporal PSD

The dependence of \(\langle P_T \rangle\) on averaging timescale \(T\) can be represented by an integral of the temporal PSD. Here we show how \(\langle P_T \rangle\) and \(\langle \hat{P}_T \rangle\) contain information about variability on different timescales. Applying Parseval’s theorem to Equation (6) as \(\tau \to \infty\), the integral transforms into the Fourier domain,
\[
\lim_{\tau \to \infty} \langle P_T \rangle = \lim_{\tau \to \infty} \frac{1}{T} \int d\omega \Phi(u, v, \omega)(1 - e^{-2\pi\tau^2\omega^2})^2. \tag{18}
\]
Subtracting a moving average on different timescales is thus equivalent to applying a high-pass filter to the temporal variability at each spatial scale. Where \(\Phi\) is a red-noise process,154 this integral is dominated by frequencies near

153 The KORAL runs used for the exploration in Section A.3 have nonuniform frame spacing. For these, each frame’s \(\Delta t\) is found as half the time difference between the previous and successive frames. For the first and last frame, the simulation-averaged frame spacing is used. No temporal Fourier transforms are taken for these simulations.

154 The Gaussian smoothing procedure we employ during averaging effectively softens the high-pass filter, to the point where there exists a maximum power-law index (equal to 4) beyond which the spectrum remains red even after filtering. The GRMHD simulations used in this work typically have power-law spectral indices less than this value in the spatial and temporal scales of interest.
Note. Collectively, \((P_T)\) and \((\hat{P}_T)\) are referred to as the variability PSDs.

\[
\omega \sim 1/2\pi T, \quad \text{so} \quad \langle P_T \rangle \approx \Phi(u, v, 1/2\pi T). \tag{19}
\]

This equivalence is demonstrated for a GRMHD simulation in Section 4.2. As \(T \to \infty\), \(\langle P_T \rangle\) is formally affected by the inability to measure infinite times but will theoretically converge to the variance map of \(F_{xy}(I)\). On finite timescales, it can be shown that

\[
d \ln \langle P_T \rangle \bigg|_T = \frac{1}{\langle P_T \rangle_T} \int d\omega \frac{d\Phi}{d \ln \omega}(1 - e^{-2\pi^2\omega^2 T^2}). \tag{20}
\]

For a red-noise process, we can thus directly relate the temporal behavior of \(\langle P_T \rangle\) to the spatiotemporal PSD,

\[
d \ln \langle P_T \rangle \bigg|_T \approx 1 - \left(\frac{d \ln \Phi}{d \ln \omega}\right)_{\omega=1/2\pi T}. \tag{21}
\]

More specifically, if \(\Phi \sim \omega^{-\eta}\) near the dominating frequencies at \(\omega \sim 1/2\pi T\), then \(\langle P_T \rangle \sim T^{\eta-1}\). An equivalent characterization exists for the light-curve-normalized version, with the replacement \(P_T \to \hat{P}_T\) and \(I \to \hat{I}\).

### 4. Application to GRMHD

We now apply the formalism presented in Section 3 to the GRMHD library, and we characterize the resulting variability PSDs. Specifically, we explore how the variability PSDs depend on the averaging timescale, the light-curve normalization, and how the spatial component depends on the orientation. For the representative examples provided in this section, we use a SANE simulation from model set A, with \(a_\bullet = 0.94\), \(R_{\text{high}} = 40\), and an inclination of 90\(^\circ\). It includes five realizations, each of length 5000 \(GM/c^3\) from 5000 \(GM/c^3\) to 30,000 \(GM/c^3\) (only the last three are listed in Table 2). \(^{155}\) Where otherwise stated, we instead use simulation set B, which samples frames at a higher cadence.

In Sections 4.1 and 4.2, we discuss the red noise of the spatiotemporal PSD and describe the equivalence between the mean subtraction procedure and applying a filter to the spatiotemporal PSD. In Section 4.3, we show the dependence on the averaging timescale. In Section 4.4, we demonstrate the effect of normalizing by the light curve. In Section 4.5, we discuss the short- and long-baseline limits, and we explore azimuthal dependence in Section 4.6.

\(^{155}\) For reference, 5000 \(GM/c^3\) is about 30 hr for Sgr A* and 5 yr for M87*.
4.1. Red-noise Power Spectrum

Figure 2 shows the three-dimensional spatiotemporal PSD of a representative GRMHD simulation, and Figure 3 shows two-dimensional and one-dimensional slices of the same PSD. A single point \((u, v, \omega)\) in \(\Phi(u, v, \omega)\) is a measure of the variability at a temporal frequency \(\omega\) and at a spatial frequency \((u, v)\). The GRMHD simulations produce characteristically red-noise variability in all dimensions, with the spatiotemporal PSD generically increasing toward small \((u, v, \omega)\) and exhibiting its maximum power at \((u, v, \omega) = (0, 0, 0)\).

4.2. Mean Subtraction

The spatiotemporal PSD evaluated at \((u, v) = (0, 0)\) represents the complex square of the Fourier transform of the light curve. The top panel of Figure 4 shows an example light curve \(L(t)\), time-smoothed versions of that light curve \(\tilde{L}(t)\) for several choices of the averaging timescale \(T\) (solid lines), and the residual light curve after subtracting out these averaged versions (dashed lines). The bottom panel of Figure 4 shows the temporal PSD of these residual light curves, equivalent to the \((u, v) = (0, 0)\) slice of the spatiotemporal PSD of the mean-subtracted image.

The act of constructing an average mean-subtracted PSD on an averaging timescale \(T\) is equivalent\(^{156}\) to multiplying \(\Phi\) by the filter described in Equation (18). Due to the red-noise nature of the system, the resulting (i.e., filtered) spatiotemporal PSD peaks near \(\omega = (2\pi T)^{-1}\); this effect is visually apparent in the bottom panel of Figure 4. One consequence of this effect is that when performing an integral of the mean-subtracted PSD with respect to \(\omega\) (such as in Equation (18)), the result will be dominated by a single timescale; we thus have \(\langle P_T \rangle(0, 0) \approx \Phi(0, 0, 1/2\pi T)\).

This equivalence between the mean subtraction procedure and a simple evaluation of the spatiotemporal PSD extends beyond just the light curve (i.e., it holds even when \(u\) or \(v\) is nonzero). For all spatial scales on which the temporal variability is a red-noise process, \(\langle P_T \rangle(u, v)\) will approximately measure \(\Phi(u, v, 1/2\pi T)\).

4.3. Spatial Properties and Averaging Timescale

The spatiotemporal PSD evaluated at \(\omega = 0\) represents the complex square of the Fourier transform of the average image (see, e.g., the red curves in the middle column of Figure 3). Because these GRMHD simulations designed for Sgr A* typically exhibit a ring of emission in the image domain (see Figure 1), they often show a Bessel function–like behavior in the Fourier domain with a first minimum near \(\sim 2–3 \text{ GM}\).

The spatiotemporal PSD evaluated at \(\omega > 0\) measures the variability power at a particular temporal frequency, as a function of the spatial size of the contributing fluctuations in the image. Figure 5 shows a one-dimensional slice of the average mean-subtracted PSD, \(\langle P_T \rangle(0, v)\), for a MAD simulation with \(a = 0.94\), \(R_{\text{high}} = 40\), and an inclination of 60° (the slowlight simulation from model set B) on several different averaging timescales \(T\). From the arguments in Section 4.2, \(\langle P_T \rangle(0, v) \approx \Phi(0, 0, 1/2\pi T)\), and thus \(\langle P_T \rangle(0, v)\) contains information about the variability on approximately a single timescale. We can see that for large \(T\) the amount of variability as measured by this average mean-subtracted PSD is larger. As \(T \to \infty\), \(\langle P_T \rangle(0, v)\) shows a broken power-law behavior, exhibiting a flat spectrum in \(v\) up to some threshold spatial frequency, and then a falling power law at higher spatial frequencies. On spatial scales larger than about 200 \(\mu\text{s}\) (corresponding to spatial frequencies smaller than 1 \(\text{GM}\)) the simulations contain no significant structure, and thus no significant variability, so the average mean-subtracted PSD on these scales must be flat. On smaller spatial scales (i.e., higher spatial frequencies), the observed power-law behavior reflects a combination of the average image structure and the dissipative MHD turbulence cascade; we explore this small-scale behavior in more detail in Section 4.5. For shorter averaging timescales, the total amount of variability power decreases and the location of the break moves to higher spatial frequencies.

The zero-baseline limit of \(\langle P_T \rangle(0, v)\) only measures properties of the light curve. Specifically, as in Equation (15), it measures the variance of the light curve on the timescale \(T\). Figure 5 shows that \(\langle P_T \rangle(0, 0)\) for this simulation roughly scales as \(T^\alpha\) for small \(T\), so the temporal power spectrum of the light curve should fall as \(\omega^{-3}\) (see Section 3.3). On large finite spatial scales, for all \(T\) where \(\langle P_T \rangle(0, v)\) is constant in \(v\), the temporal power spectrum of the variability should scale with \(\omega\) identically to the light curve. We note that other simulations (i.e., with different accretion parameters) can in general have different temporal spectral power-law indices, but this falling power-law behavior at large \(\omega\) is common to all.

\(^{156}\) This equivalence only strictly holds in the limit where the total duration of the light curve is much longer than the averaging timescale \(T\). Effects when this property does not hold are explored in Section A.3.
4.4. Effects of Normalization by the Light Curve

On short baselines, we expect the PSD to be dominated by variability that is heavily correlated with the light-curve variability. By light-curve-normalizing as described in Section 3.1, we thus expect to remove a large fraction of the short-baseline variability. The two rows in Figure 3 show the same simulation before and after normalizing by the light curve. By definition, all variability is removed from \((u, v) = (0, 0)\), but we further see a decrease in the PSD at all spatial frequencies below \(\sim 1\) G\(\lambda\). We interpret this decrease in power on small spatial frequencies to mean that the majority of variability on large spatial scales is indeed strongly correlated with the variability in the light curve. At higher spatial frequencies, the PSD remains unchanged, indicating that the variability on small spatial scales is independent of the large-scale variability.

Figure 6 shows the dependence of \(\langle \hat{P}_T \rangle\) on the averaging timescale \(T\). For spatial frequencies larger than the break location, \(\langle \hat{P}_T \rangle\) shows the same behavior as \(\langle P_T \rangle\), i.e., a power law with the same index. The break locations of the two PSDs are not equivalent, with \(\langle \hat{P}_T \rangle\) breaking at smaller spatial scales than \(\langle P_T \rangle\). On larger spatial scales, the flat part of \(\langle P_T \rangle\) has been reduced to an increasing power law in \(v\) with a power-law index that is independent of \(T\).

4.5. Short- and Long-baseline Power Law

Figure 7 shows \(\langle P_T \rangle(0, v)\) and \(\langle \hat{P}_T \rangle(0, v)\) for the same simulation as used in Section 4.1, as well as the PSD of the average image. On the smallest spatial frequencies, we can see the \(\langle \hat{P}_T \rangle(0, v) \propto v^2\) behavior expected from Section 3.2, and this behavior continues up to a spectral break near \(\sim 1\) G\(\lambda\). The higher-order terms from Equation (14) must therefore remain subdominant to the \(v^2\) term below the break. From this empirical behavior, we can deduce that the location of the break lies approximately on the \(v^2\) power law defined in Equation (16). The break therefore contains information about the variability of the centroid of emission in the total image.

For any image structure \(I(x, y, t)\), we can generically decouple the time-variable component from the average image as

\[
I(x, y, t) = \bar{I}(x, y) + \delta I(x, y, t) = \bar{I}(x, y)[1 + \delta(x, y, t)],
\]

where \(\bar{I}(x, y)\) is understood to be the time-averaged image as \(T \to \infty\) and we have defined \(\delta \equiv \delta I/I\). Upon applying our mean-subtracting procedure (see Equation (6)), we would measure
where the symbol represents convolution. If the Fourier transforms, \( \mathcal{F}_{xy} \) and \( \mathcal{F}_{xy} \), have uncorrelated phases, and if both have amplitudes that fall as a power law at large \((u, v)\), then \( P_T(u, v, t) \) will be dominated by the shallower power-law component at long baselines.

Figure 7 shows that \( \langle P_T \rangle \) follows a similar power law at large spatial frequencies to the Fourier transform of the average image, suggesting that the variability in GRMHD simulations is perhaps better described by an uncorrelated multiplicative random field \( \delta \), rather than an uncorrelated additive field \( \delta I \). Under this assumption, we would expect the long-baseline power law of the variability in GRMHD simulations to match that of their average images, and we would conclude that the spectrum of the “true” variability \( \delta \) is steeper than that of the average image. Alternatively, the average image power law could be steeper, in which case the long-baseline behavior of the variability PSDs would be dominated by \( \delta \).

At the longest baselines, there is a slight upturn in the squared amplitudes of the average image and in both variability PSDs. We attribute this to the limited numerical accuracy with which image data from the GRMHD image library are saved (keeping only six nonzero digits). The associated truncation error introduces a white-noise floor that, while negligible for the baselines probed by the EHT, can dominate above 100 G\( \lambda \).

4.6. Azimuthal Dependence

Figure 8 shows \( \langle \hat{P}_T \rangle (u, v) \) as \( T \to \infty \) for the same simulation as in Figures 2 and 3. Near \((u, v) = (0, 0)\), the average residual PSD has little power owing to the light-curve normalization. On short baselines, from Equation (16), a constant \( \langle \hat{P}_T \rangle (u, v) \) occurs on an elliptical contour, with the minor and major axis in the direction of the largest and smallest variation in the centroid of emission, respectively. This elliptical behavior...
remains as long as the $u^2$ dependence, up until the break location. The major and minor axes are often either aligned with or perpendicular to the direction of the black hole spin axis (see Section C). On spatial frequencies larger than the break location, the power law falls with an index determined by the average image in every direction (see Section 4.5).

5. Variability for Different Accretion Flows

Given our understanding of the morphology of the GRMHD variability PSDs from Section 4, we now explore the trends seen with different simulation types.

5.1. Universal Power Law

Figure 9 shows the visibility amplitudes of the average image, $\langle P_T \rangle$, and $\langle \hat{P}_T \rangle$ for all the simulations in model set A for infinite averaging time $T$. We explore these PSDs only along and orthogonal to the direction of the black hole spin axis, as these typically encompass the extrema of the variability (see Section 4.6 and Appendix C). The variability for each simulation follows the behavior described in Section 4, and each variability PSD can be approximated as a broken power-law. Moreover, the different simulations are remarkably consistent over different magnetic field configurations, black hole spins, temperature prescriptions, and viewing inclinations, with similar broken power-law parameters.

The short-baseline behavior of $\langle P_T \rangle$ and $\langle \hat{P}_T \rangle$ that rises as $|u|^2$ match the expectations from a finite size geometry. The amount of power at short baselines and the location of the break are unique to each simulation and likely contain physical information about the underlying variability. Between about 2 and 30 $G\lambda$, the long-baseline behavior is similar to that of the underlying average image. This suggests a common origin for both, e.g., the variability could be described as an average image multiplied by an uncorrelated random field (see Section 4.5). This behavior can occur if the local variability in the fluid is random and caused by processes that scale with the density (e.g., rotational velocity, shearing). The global variability would then scale with the average density, and where the emitted intensity scales with the density, the variability in the observed image would appear multiplicative to the average image. However, the departure between the variability PSDs and that of the background average image above 30 $G\lambda$ implies that the relationship between the small-scale fluctuations and the time-averaged image structure is more complicated. More precise identification of the physical source of the universality of the variability PSDs is complicated by the interactions of the global accretion state, the emissivity, and the radiative transfer along photon trajectories, and we leave its study to future work.

5.2. PSD Dependence on Flow Parameters

We now turn our attention to understanding the dependence of the variability PSDs on GRMHD simulation parameters. The
parameter space and possible PSD measurements are vast, and trends across simulations are typically limited to small pieces of the parameter space. Therefore, in this section we limit ourselves to explaining the trends across models for two measurements from the variability PSDs and briefly comment on other extensions.

Figure 10 shows the measurement of \( \langle \hat{P}_T \rangle \) for the same simulations as in Figure 2 (i.e., a SANE simulation from model set A with \( \alpha_s = 0.94 \), \( R_{\text{high}} = 40 \), and \( i = 90^\circ \)). The top panel shows \( \langle \hat{P}_T \rangle \) in the infinite \( T \) limit, with zero power at the center corresponding to the light-curve normalization. The bottom panel shows a horizontal and vertical slice.

MAD simulations show little change across spin, inclination, or \( R_{\text{high}} \). This is not necessarily surprising: in MAD flows, the emission is typically produced close to the event horizon and is centered around the midplane, in the region where filamentary accretion strands are separated by the strong magnetic pressures and plunge from intermediate radius down to the horizon. When the majority of emission comes from this violent, disorganized region, the precise details of the accretion model and thermodynamics may not strongly influence the detailed morphology of the emission.

For SANE simulations, however, there is a clear decrease in variability for larger corotating spin, larger \( R_{\text{high}} \), and more edge-on viewer inclination. In contrast with the MADs, the morphology of the SANE emission region can be much more strongly affected by the details of the model, and the split between \( R_{\text{high}} = 10 \) and \( R_{\text{high}} = 40 \) coincides with a change from emission dominated by

---

**Figure 8.** Azimuthal dependence of \( \langle \hat{P}_T \rangle \) for the same simulations as in Figure 2 (i.e., a SANE simulation from model set A with \( \alpha_s = 0.94 \), \( R_{\text{high}} = 40 \), and \( i = 90^\circ \)). The top panel shows \( \langle \hat{P}_T \rangle \) in the infinite \( T \) limit, with zero power at the center corresponding to the light-curve normalization. The bottom panel shows a horizontal and vertical slice.

**Figure 9.** The average image visibilities (top), average mean-subtracted PSD (middle), and average residual PSD (bottom) along the \( u \) (left) and \( v \) (right) directions for the 360 simulations in model set A (KHARMA). All PSDs use infinite averaging time. For \( \langle \hat{P}_T \rangle \), all simulations show a flat spectrum for small baselines and a power-law drop-off at large baselines. For \( \langle \hat{P}_T \rangle \), all simulations show a broken power-law spectrum, rising as \( |u|^2 \) for small baselines, and a power-law drop-off at large baselines. The long-baseline power laws are remarkably consistent among different simulation types.
the disk or the jet sheath, respectively. For disk-dominated $R_{\text{high}} \gtrsim 10$, the variability is much more pronounced for face-on inclinations, where any part of the disk may light up, than for edge-on inclinations, where the emission, and therefore variability, is limited to the Doppler-boosted gas moving toward the observer on only one side of the black hole. For jet-dominated $R_{\text{high}} \gtrsim 40$, the emission comes from further off the midplane and is not as beamed, thus changing inclination has little effect on the emission region size orthogonal to the spin axis. Note that for these jet-dominated models the component along the spin axis does depend on inclination.

Figure 11 is analogous to Figure 10, but showing $\langle \hat{P}_r \rangle (0, 0.5 \, \lambda \alpha, 0) / \langle \hat{P}_r \rangle (0.5 \, \lambda \alpha, 0)$ for infinite average time $T$. This is in the short-baseline regime, and therefore we can interpret trends through the axial ratio of the covariance ellipse of the distribution of centroids of emission. MAD and SANE simulations are split, with MADs seeing more variability along the disk where the average bulk of the emission is concentrated, and SANEs seeing more along the jet, which is especially lit up in models with large $R_{\text{high}}$. SANE simulations show a split between disk- and jet-dominated emission as in Figure 10. The simulations dominated by jet emission have small ratios (i.e., $<1$), since the variability comes from different parts of the jet and counterjet lighting up. This effect becomes more pronounced at edge-on inclinations where the average image resembles two disjoint emitting regions along the direction of the spin axis.

For long baselines, the variability PSDs are related to the power spectrum of the average image. Therefore, any trends seen in the variability would likely be similar to trends in the average image, which can more easily be obtained from data. On short timescales (small $T$), simulations show a clear power-law dependence on $T$ with a temporal break, which, like the two PSD measurements
discussed in this section, can discriminate accretion flow types. However, on shorter timescales, it becomes increasingly difficult to provide a physical origin for the PSD measurement trends, and we leave its study to future work.

5.3. Application to Sgr A*

Current EHT observations of Sgr A* provide only sparse temporal and spatial coverage of the \((\mu, \nu)\)-domain. A measurement of the variability PSDs from EHT data is further obscured by instrumental and calibration effects. However, tests on synthetic data from GRMHD simulations in A. E. Broderick et al. (2022) show that \((\hat{P}_T)\) can be well constrained with current observations for \(2 \, \text{G}\lambda < |\mu| < 6 \, \text{G}\lambda\).

Motivated by these findings, we show measurements of \((\hat{P}_T)\) from GRMHD simulations that have been generated to mimic the EHT observations of Sgr A*. For each simulation, we diffractively scatter the image with the kernel from Johnson et al. (2018) to include the effects of interstellar scattering. We then average its variability PSDs over all possible position angles of the scattering screen, and then we azimuthally average. Finally, we fit each resulting scattered \((\hat{P}_T)_{0}\) with a power law between \(2 \, \text{G}\lambda < |\mu| < 6 \, \text{G}\lambda\) and take our measurement to be the value of the fit at \(|\mu| = 4 \, \text{G}\lambda\). This procedure is shown for the simulation used in Section 4.1 in Figure 12.

The equivalent of the full scattered \((\hat{P}_T)_{0}\) is measured from EHT data in Paper IV, and the amplitude and slope of this power-law measurement are compared to data and interpreted in Paper V.\(^{157}\)

Figure 13 shows this measurement constructed for infinite averaging time \(T\). Unlike at short baselines, \((\hat{P}_T)\) near \(4 \, \text{G}\lambda\) is similar for many of the considered accretion flow parameters, predicting a universal amount of GRMHD variability to within an order of magnitude.

The main trend is with inclination, with face-on models showing less variability than edge-on models. This is likely mirroring average visibilities, where face-on models have a larger area of emission and thus larger variability. Further discriminatory power is limited by the noisiness of the simulations, with separate windows and codes providing the leading uncertainty in the measurement of \((\hat{P}_T)\).

6. Summary and Conclusions

Motivated by the EHT observations of Sgr A*, in this paper we have developed a framework for characterizing the spatiotemporal power spectrum of variability observed toward an accreting supermassive black hole system. We have applied this framework to the library of GRMHD simulations and associated GRRT images produced by Paper V, which span a range of physical properties (magnetic flux configurations, black hole spin, temperature prescription), numerical properties (code, resolutions, initialization, simulation length, time from initialization, and other approximations), and observer locations. While these models were produced specifically to match the expected emission from Sgr A* during the 2017 EHT observation period, one can reasonably expect our findings to be applicable to a range of systems in the optically thin regime, following a proper scaling of a black hole mass and total flux. We find that the variability power spectrum is generically a red-noise process in both the temporal and spatial dimensions, i.e., the simulations exhibit the highest variability power on the longest timescales and on the largest spatial scales. To focus on variability in excess of the spatiotemporal mean, we have considered primarily the PSD behavior after subtracting out the time-averaged source structure and normalizing by the light curve.

All of the GRMHD simulations show remarkably similar structure in their variability PSDs. The spatial PSDs along any direction can be described by a broken power law, with a small range of long-baseline indices (\(\sim 2–3\)) and break locations (\(\sim 1–3 \, \text{G}\lambda\)). On long baselines, the power-law index of the variability PSD tracks that of the Fourier transform of the average image, which is itself similar across simulations. The observed universality in the long-baseline PSD behavior could arise if the variability came in the form of an average image whose structure is modulated by an uncorrelated, multiplicative random field. In this picture, it is the Fourier transforms of the average images that are similar among simulations, and the random field could be caused by any process (e.g., MHD turbulence) that has a long-base power-law index steeper than that of the average image.

On any given spatial scale, the power in the variability about the average image increases with averaging time up to some breaking timescale, beyond which it remains constant. That is, for every spatial scale there exists a maximum necessary averaging time, above which the variability power on that spatial scale remains unchanged with increased averaging. On short baselines (below \(\sim 1 \, \text{G}\lambda\)), this breaking timescale is greater than \(10^6 \, \text{GM} / \mu^2\), and it is shorter on longer baselines. The implication for EHT observations of M87* is that the variability “noise” seen across a single night of observing should be approximately constant with baseline length, and the expectation from GRMHD simulations is that the average

\(^{157}\) The convention in Paper V differs by a factor of 2, due to the use of visibility amplitudes instead of complex visibilities. The rest of the procedure is identical.
magnitude of this variability should be less than 1% of the total flux. Across a week of observations, the expected variability power increases to several percent, though it remains flat with baseline length on the range of baselines probed by the EHT. These expectations qualitatively match the observed behavior of M87*, which the EHT found to exhibit little or no intraday variability and a modest amount of interday variability (M87 Paper IV). Significant structural variation can be expected on timescales of months and years, which has been observationally confirmed by Wielgus et al. (2020). For Sgr A*, we expect that the variability should manifest as a broken power-law “noise,” with large contributions on short baselines and decreasing contributions on longer baselines. Normalizing the visibility amplitudes by the light curve removes the largest component of this variability noise.

The azimuthal structure of the variability PSDs tracks the average images, that is, the variability typically occurs where the emission does. Especially on longer baselines where the variability PSDs and average image are similar, we expect the variability to discriminate between accretion flows less than the average images could, and the latter are observationally easier to measure. The variability can better distinguish different accretion flows from measurements on short baselines (|u| ≲ 1 Gλ), where the variability is larger. Here the variability power is proportional to the variance in the centroid of emission, the most variable spatial mode after the light curve. The entire temporal PSD of a short-baseline measurement of the variability (i.e., for a time series of the centroid of emission) for timescales greater than ∼30 GM/c² contains unique discriminating power. However, with the current EHT instrumental effects, such a measurement would contain degeneracies with the average image.

We find that a measurement of the average residual PSD on short baselines—and thus a measurement of the covariance ellipse of the centroid of emission—wields some power to discriminate between the different accretion flows in our GRMHD library. In particular, the MAD and SANE accretion states exhibit distinct emission centroid behaviors relative to the direction of the black hole spin axis. A centroid ellipse that is strongly elongated along the direction of the spin axis suggests a variable jet, which is often seen in SANE simulations that support hot funnels. In such a case, the magnitude of variability orthogonal to the spin axis can constrain the black hole spin, with more positive spins producing less variability in this direction. Centroid ellipses that are strongly elongated along the direction of the disk (i.e., assumed perpendicular to the spin axis) are instead more typical of edge-on MAD states, which have more equatorial emission structures. In this case, spin constraints are more difficult because the variability in MAD simulations does not significantly change with spin or temperature prescription. Circular centroid ellipses suggest a face-on accretion flow, or perhaps an emission region that is farther from the black hole. In such cases, the amount of variability in the direction of the disk could distinguish whether the accretion flow is in the SANE state with emission dominated by hot electrons in the disk. For each of the cases described, the temporal behavior also holds significant discriminating power, though it remains a more uncertain measurement from the simulations used here and is left for future study.

Lastly, we note that the work and conclusions rely entirely on the assumption that the accretion flows around black holes and Sgr A* in particular are described by ideal GRMHD and that the assumptions within those simulations are valid. We have attempted to address the latter by individually studying small excursions in resolution, the fastlight approximation, accretion flow evolution, random initialization, and further assumptions made with different simulation codes. Primarily, we find that the variability evolves on long timescales and that even simulations run for as long as 30,000 GM/c³ may be insufficient for the characterization of variability statistics.
(although the opinions expressed in this work are those of the author(s) and do not necessarily reflect the views of these foundations); NASA Hubble Fellowship grant HST-HF2-51431.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555; the Academy of Finland (projects 274477, 284495, 312496, 315721); the Agencia Nacional de Investigación y Desarrollo (ANID), Chile via NCN19_058 (TITANs) and Fondecyt 1221421, the Alexander von Humboldt Stiftung; an Alfred P. Sloan Research Fellowship; Allegro, the European ALMA Regional Centre node in the Netherlands, the NL astronomy research network NOVA and the astronomy institutes of the University of Amsterdam, Leiden University, and Radboud University; the Institute for Advanced Study; the China Scholarship Council; Consejo Nacional de Ciencia y Tecnología (CONACYT, Mexico, projects U0004-246083, U0004-259839, F0003-272050, M0037-279006, F0003-281692, 104497, 275201, 263356); the Delaney Family via the Delaney Family John A. Wheeler Chair at Perimeter Institute; Dirección General de Asuntos del Personal Académico-Universidad Nacional Autónoma de México (DGAPA-UNAM, projects IN112417 and IN112820); the European Research Council Synergy Grant “BlackHoleCam: Imaging the Event Horizon of Black Holes” (grant 610058); the Generalitat Valenciana postdoctoral grant APOSTD/2018/177 and GenT Program (project CIDEGENT/2018/021); MICINN Research Project PID2019-108995GB-C22; the European Research Council for advanced grant “JETSET: Launching, propagation and emission of relativistic jets from binary mergers and across mass scales” (grant No. 884631); the Istituto Nazionale di Fisica Nucleare (INFN) sezione di Napoli, iniziative specifiche TEONGRAV; the two Dutch National Supercomputers, Cartesius and Snellius (NWO grant 2021.013); the International Max Planck Research School for Astronomy and Astrophysics at the Universities of Bonn and Cologne; DFG research grant “Jet physics on horizon scales and beyond” (grant No. FR 4069/2-1); Joint Princeton/Flatiron and Joint Columbia/Flatiron Postdoctoral Fellowships, with research at the Flatiron Institute supported by the Simons Foundation; the Japanese Government (Monbukagakusho: MEXT) Scholarship; the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for JSPS Research Fellowship (JP17J08829); the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS, grants QYZDJ-SSW-SLH057, QYZDJ-SSW-SYS008, ZDBS-LY-SLH011); the Leverhulme Trust Early Career Research Fellowship; the Max-Planck-Gesellschaft (MPG); the Max Planck Partner Group of the MPG and the CAS; the MEXT/ JSPS KAKENHI (grants 18KK0090, JP21H01137, JP18H03721, JP18K13594, 18K03709, JP19K14761, 18H01245, 25120007); the Malaysian Fundamental Research Grant Scheme (FRGS) FRGS/1/2019/STG02/UM/02/6; the MIT International Science and Technology Initiatives (MISTI) Funds; the Ministry of Science and Technology (MOST) of Taiwan (103-2191-M-001-010-MY2, 105-2112-M-001-025-MY3, 105-2119-M-001-042, 106-2112-M-001-011, 106-2119-M-001-013, 106-2119-M-001-027, 106-2923-M-001-005, 107-2119-M-001-017, 107-2119-M-001-020, 107-2119-M-001-041, 107-2119-M-110-005, 107-2923- M-001-009, 108-2112-M-001-048, 108-2112-M-001-051, 108-2923-M-001-002, 109-2112-M-001-025, 109-2124-M-001-005, 109-2923-M-001-001, 110-2112-M-003-007-MY2, 110-2112-M-001-033, 110-2124-M-001-007, and 110-2923-M-001-001); the Ministry of Education (MoE) of Taiwan Yushan Young Scholar Program; the Physics Division, National Center for Theoretical Sciences of Taiwan; the National Aeronautics and Space Administration (NASA, Fermi Guest Investigator grant 80NSSC20K1567, NASA Astrophysics Theory Program grant 80NSSC20K0527, NASA NuSTAR award 80NSSC20K0645); the National Institute of Natural Sciences (NINS) of Japan; the National Key Research and Development Program of China (grant 2016YFA0400704, 2017YFA0402703, 2016YFA0400702); the National Science Foundation (NSF, grants AST-0096454, AST-0352953, AST-0521233, AST-0705062, AST-0905844, AST-0922984, AST-1126433, AST-1140030, DGE-1144085, AST-1207704, AST-1207730, AST-1207752, MRI-1228509, OPP-1248097, AST-1310896, AST-1555365, AST-1614868, AST-1615796, AST-1715061, AST-1716327, AST-2034306); the Natural Science Foundation of China (grants 11650104472, 10625314, 11721303, 11725312, 11873028, 11933007, 11991052, 11991053, 12192220, 12192223); NWO grant No. OCENW.KLEIN.113; a fellowship of China Postdoctoral Science Foundation (2020M671266); the Natural Sciences and Engineering Research Council of Canada (NSERC, including a Discovery Grant and the NSERC Alexander Graham Bell Canada Graduate Scholarships-Doctoral Program); the National Youth Thousand Talents Program of China; the National Research Foundation of Korea (the Global PhD Fellowship Grant: grants NRF-2015H1A20137352, the Korea Research Fellowship Program: NRF-2015H1D3A1066561, Basic Research Support grant 2019RF1A1059721, 2022R1C1C1005255); the Dutch Organization for Scientific Research (NWO) VICI award (grant 639.043.513) and Spinoza Prize SPI 78-409; and the YCAA Prize Postdoctoral Fellowship. L.M. gratefully acknowledges support from an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award no. AST-1903847. T.K. is supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) as a “Program for Promoting Researches on the Supercomputer Fugaku” (Toward a unified view of the universe: from large-scale structures to planets, JPMXP01020200109) and JICFuS. R.P.D. and I.N. acknowledge funding by the South African Research Chairs Initiative, through the South African Radio Astronomy Observatory (SARAO, grant ID 77948), which is a facility of the National Research Foundation (NRF), an agency of the Department of Science and Innovation (DSI) of South Africa.

We thank the Onsala Space Observatory (OSO) national infrastructure, for the provisioning of its facilities/observational support (OSO receives funding through the Swedish Research Council under grant 2017-00648); the Perimeter Institute for Theoretical Physics (research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research, Innovation and Science); the Spanish Ministerio de Ciencia e Innovación (grants PGC2018-098915-B-C21, AYA2016-80889-P, PID2019-108995GB-C21, PID2020-117404GB-C21); the University of Pretoria for financial aid in the provision of the new Cluster Server nodes and SuperMicro (USA) for a SEEDING GRANT approved toward these nodes in 2020; the State Agency for Research of the Spanish MCIU through the “Center of Excellence Severo Ochoa” award for the Instituto de Astrofísica de Andalucía (SEV-2017-0709); the Toray Science Foundation; the Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía (grant P18-FR-1769), the Consejo...
Superior de Investigaciones Científicas (grant 2019AEP112); the M2FINDERS project, which has received funding by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (grant agreement No. 101018682); the US Department of Energy (USDOE) through the Los Alamos National Laboratory (operated by Triad National Security, LLC, for the National Nuclear Security Administration of the USDOE (contract 89233218CNA000001)); the European Union’s Horizon 2020 research and innovation program under grant agreement No. 730562 RadioNet; Shanghai Pilot Program for Basic Research, Chinese Academy of Science, Shanghai Branch (JCYJ-SHFY-2021-013); ALMA North America Development Fund; the Academia Sinica; Chandra DD7-18089X and TM6-17006X; and the GentI Program (Generalitat Valenciana) Project CIDEGENT/2018/021. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), supported by NSF grant ACI-1548562, and CyVerse, supported by NSF grants DBI-0735191, DBI-1265383, and DBI-1743442. The XSEDE Stampede2 resource at TACC was allocated through TG-AST170024 and TG-AST080026N. The XSEDE JetStream resource at PTI and TACC was allocated through AST170028.

The simulations were performed in part on the SuperMUC cluster at the LRZ in Garching, on the LOEWE cluster in CSC in Frankfurt, and on the HazelHen cluster at the HLRS in Stuttgart. This research was enabled in part by support provided by Compute Ontario (http://computeontario.ca), Calcul Quebec (http://www.calculquebec.ca), and Compute Canada (http://www.computecanada.ca). C.C. acknowledges support from the Swedish Research Council (VR).

We thank the staff at the participating observatories, correlation centers, and institutions for their enthusiastic support. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2016.1.01154.V. ALMA is a partnership of the European Southern Observatory (ESO; Europe, representing its member states), NSF, and National Institutes of Natural Sciences of Japan, together with National Research Council (Canada), Ministry of Science and Technology (MOST; Taiwan), Academia Sinica Institute of Astronomy and Astrophysics (ASIAA; Taiwan), and Korea Astronomy and Space Science Institute (KASI; Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, Associated Universities, Inc. (AUI)/NRAO, and the National Astronomical Observatory of Japan (NAOJ). The NRAO is a facility of the NSF operated under cooperative agreement by AUI. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract No. DE-AC05-00OR22725. We also thank the Center for Computational Astrophysics, National Astronomical Observatory of Japan.

Support for this work was also provided by the NASA Hubble Fellowship grant HST-HF2-51431.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. H.O. and G.M. were supported by Virtual Institute of Accretion (VIA) postdoctoral fellowships from the Netherlands Research School for Astronomy (NOVA). APEX is a collaboration between the Max-Planck-Institut für Radioastronomie (Germany), ESO, and the Onsala Space Observatory (Sweden). The SMA is a joint project between the SAO and ASIAA and is funded by the Smithsonian Institution and the Academia Sinica. The JCMT is operated by the East Asian Observatory on behalf of the NAOJ, ASIAA, and KASI, as well as the Ministry of Finance of China, Chinese Academy of Sciences, and the National Key Research and Development Program (No. 2017YFA0402700) of China and Natural Science Foundation of China grant 11873028. Additional funding support for the JCMT is provided by the Science and Technologies Facility Council (UK) and participating universities in the UK and Canada. Simulations were performed in part on the SuperMUC cluster at the LRZ in Garching, on the LOEWE cluster in CSC in Frankfurt, on the HazelHen cluster at the HLRS in Stuttgart, and on the P2.0 and Siyuan Mark-I at Shanghai Jiao Tong University. The computer resources of the Finnish IT Center for Science (CSC) and the Finnish Computing Competence Infrastructure (FCCI) project are acknowledged. J.O. was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1A6A3A01086420; 2022R1C1C1005255). We thank Martin Shepherd for the addition of extra features in the Difmap software that were used for the CLEAN imaging results presented in this paper. The computing cluster of Shanghai VLB1 correlator supported by the Special Fund for Astronomy from the Ministry of Finance in China is acknowledged. This work was supported by the Brain Pool Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (019H1D3A1A01102564). This research is part of the Frontera computing project at the Texas Advanced Computing Center through the Frontera Large-Scale Community Partnerships allocation AST20023. Frontera is made possible by National Science Foundation award OAC-1818253. This research was carried out using resources provided by the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy Office of Science.

The LMT is a project operated by the Instituto Nacional de Astronomía, Óptica, y Electrónica (Mexico) and the University of Massachusetts at Amherst (USA). The IRAM 30 m telescope on Pico Veleta, Spain, is operated by IRAM and supported by CNRS (Centre National de la Recherche Scientifique, France), MPG (Max-Planck-Gesellschaft, Germany), and IGN (Instituto Geográfico Nacional, Spain). The SMT is operated by the Arizona Radio Observatory, a part of the Steward Observatory of the University of Arizona, with financial support of operations from the State of Arizona and financial support for instrumentation development from the NSF. Support for SPT participation in the EHT is provided by the National Science Foundation through award OPP-1852617 to the University of Chicago. Partial support is also provided by the Kavli Institute of Cosmological Physics at the University of Chicago. The SPT hydrogen maser was provided on loan from the GLT, courtesy of ASIAA. Support for this work was provided by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51494.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Jongho Park acknowledges financial support through the EACOA Fellowship awarded by the East Asia Core Observatories Association, which consists of the Academia Sinica Institute of Astronomy and Astrophysics, the National Astronomical Observatory of Japan, Center for Astronomy (ASIAA), and the University of Massachusetts at Amherst (USA). The IRAM 30 m telescope on Pico Veleta, Spain, is operated by IRAM and supported by CNRS (Centre National de la Recherche Scientifique, France), MPG (Max-Planck-Gesellschaft, Germany), and IGN (Instituto Geográfico Nacional, Spain). The SMT is operated by the Arizona Radio Observatory, a part of the Steward Observatory of the University of Arizona, with financial support of operations from the State of Arizona and financial support for instrumentation development from the NSF. Support for SPT participation in the EHT is provided by the National Science Foundation through award OPP-1852617 to the University of Chicago. Partial support is also provided by the Kavli Institute of Cosmological Physics at the University of Chicago. The SPT hydrogen maser was provided on loan from the GLT, courtesy of ASIAA. Support for this work was provided by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51494.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Jongho Park acknowledges financial support through the EACOA Fellowship awarded by the East Asia Core Observatories Association, which consists of the Academia Sinica Institute of Astronomy and Astrophysics, the National Astronomical Observatory of Japan, Center for Astronomy (ASIAA), and the University of Massachusetts at Amherst (USA). The IRAM 30 m telescope on Pico Veleta, Spain, is operated by IRAM and supported by CNRS (Centre National de la Recherche Scientifique, France), MPG (Max-Planck-Gesellschaft, Germany), and IGN (Instituto Geográfico Nacional, Spain). The SMT is operated by the Arizona Radio Observatory, a part of the Steward Observatory of the University of Arizona, with financial support of operations from the State of Arizona and financial support for instrumentation development from the NSF. Support for SPT participation in the EHT is provided by the National Science Foundation through award OPP-1852617 to the University of Chicago. Partial support is also provided by the Kavli Institute of Cosmological Physics at the University of Chicago. The SPT hydrogen maser was provided on loan from the GLT, courtesy of ASIAA. Support for this work was provided by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51494.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Jongho Park acknowledges financial support through the EACOA Fellowship awarded by the East Asia Core Observatories Association, which consists of the Academia Sinica Institute of Astronomy and Astrophysics, the National Astronomical Observatory of Japan, Center for
Astronomical Mega-Science, Chinese Academy of Sciences, and the Korea Astronomy and Space Science Institute.

The EHTC has received generous donations of FPGA chips from Xilinx Inc., under the Xilinx University Program. The EHTC has benefited from technology shared under open-source license by the Collaboration for Astronomy Signal Processing and Electronics Research (CASPER). The EHT project is grateful to T4Science and Microsemi for their assistance with Hydrogen Masers. This research has made use of NASA’s Astrophysics Data System. We gratefully acknowledge the support provided by the extended staff of the ALMA, both from the inception of the ALMA Phasing Project through the observational campaigns of 2017 and 2018. We would like to thank A. Deller and W. Brisken for EHT-specific support with the use of DiFX. We acknowledge the significance that Maunakea, where the SMA and JCMT EHT stations are located, has for the indigenous Hawaiian people.

Software: BHAC, BHSS, eht-imaging, iharm3d, ipole, KCHARMA, KORAL.

Appendix A

The Impact of GRMHD Numerical Approximations

A.1. Dependence on the Fastlight Approximation

The model sets used in this work primarily use the fastlight approximation, where the fluid output is saved at constant values of the Kerr–Schild timelike coordinate and each image is constructed using only one of these fluid outputs, effectively assuming that the background flow is constant in time as light propagates through all space. In contrast, slowlight (or finite-speed light) methods account for the evolution of the fluid by loading the changing fluid state (i.e., a sequence of constant-Kerr–Schild time snapshots) as light propagates through space. In our approach, the fluid data at an arbitrary spacetime event are synthesized by performing linear interpolation in both space (between neighboring grid zones) and time (between sequential fluid snapshots). Since the fluid velocity at the emission region has a significant component in each fluid output and the scale of the system is not much smaller than its light-crossing time, we expect an approximated fastlight image to deviate from the truth for temporal and spatial scales comparable to the gravitational scales.

To test the numerical effects of this approximation on the power spectra, we imaged a MAD simulation with $a_\text{o} = 0.94$ at a short cadence of $0.5 \, GM/c^2$ (model set B). For the fastlight imaging, each fluid output creates one image for a viewer far from the black hole. For the slowlight imaging, the same viewing screen was created in time corresponding to the time the light ray would need to travel from the black hole, so that the fastlight and slowlight images would roughly span the same accretion history. To find this offset more exactly, we compared the light curves of the two simulations and maximized their correlations relative to a time shift. The light curves are dominated by the longest timescales and thus are nearly identical to the appropriate time shift. After this calibration procedure, the two sets of images have the same cadence and nearly the same start and end points.

Figure 14 shows the effect of using the fastlight approximation on $\langle P_T \rangle$ and $\langle P_T \rangle$. Solid lines show the “true” slowlight quantities, and dashed lines show the fastlight approximation. For long timescales, the approximation holds, likely due to long-timescale variability caused by gas farther from the black hole, which only has a small component of its velocity orthogonal to the timelike coordinate. On shorter timescales, the fastlight approximation consistently suppresses variability on all spatial scales, though the magnitude of the difference is not a particularly large one.

A.2. Dependence on Simulation Resolution

Numerical GRMHD simulations lay out a grid upon which the fluid is evolved. Therefore, they set a lower boundary on the spatial size of fluctuations of the fluid properties. It is not known whether physics at the unmodeled small scales could cause large-scale deviations in the fluid properties through perhaps a different saturation of the MHD turbulence. These potential different states could manifest as different variability power spectra.

In order to test the effects of finite spatial resolution, we have performed GRMHD simulations of a MAD model with $a_\text{o} = 0.94$ in three different resolutions, 96$^3$, 128$^3$, and 192$^3$ starting from an identical initial condition (model set E).
greater resolutions and is thus outside the scope of this work.

Figure 15 shows the variability PSDs for these three simulations. Although there are differences in the variability PSDs, there is no clear trend with resolution, hinting that the main source of the variability stems from physics on the resolved larger scales. This test does not preclude the possibility of the simulations not being sufficiently resolved, but resolving such effects likely requires far greater resolutions and is thus outside the scope of this work.

A.3. Dependence on Turbulent Realization

For each of the simulations considered in this work, the total length of time available ($\tau$ in Section 3) is finite, and thus we only have access to one possible realization of turbulence. Therefore, for a given simulation length, our PSD measurement will be one random draw from a distribution of possible PSDs. Here we try to characterize any bias and the distribution width expected for such a measurement.

For the conclusions in this work, we primarily use simulations of length 5000 $GM/c^3$. As a representative, we choose a simulation in the MAD state from model set A with $a_a = -0.94$, $R_{high} = 10$, and $i = 10^5$. It contains three independent realizations. As the “true” model, we take a simulation in the MAD state from model set C, with $a_a = -0.9$, $R_{high} = 20$, and $i = 10^3$. Using only frames with cadence 50 $GM/c^3$, it covers 70,000 $GM/c^3$, thus containing the equivalent of 14 separate realizations of the representative simulations.

From each simulation, we constructed a sliding window of some shorter time. From each sliding window we calculate $\langle P_T \rangle(0,0)$ for different sliding window lengths. For each sliding window length, the error bars and points show the 16th, 50th, and 84th percentiles over every sliding window. Note that these sliding windows are not independent, and the number of uncorrelated sliding windows is approximately the ratio of the sliding window length to $\tau$. Thus, the measurements of the percentiles for ratios $\lesssim 10$ are not robust. Where there are enough independent sliding windows, these percentiles represent a distribution of potential measurements of $\langle P_T \rangle(0,0)$ for a simulation with shorter $\tau$.

Building confidence, a measurement of the median of $\langle P_T \rangle(0,0)$ using the shorter simulation does fall within the measured distribution from the longer simulation, despite the two simulations having slightly different values of spin and $R_{high}$. For both simulations, this median is constant over sliding window length, with a deviation occurring when the window length is comparable to the averaging time, $T$. This is simply because the numerical Gaussian mean subtraction from Equation (8) truncates when the simulation ends, thus shortening the effective $T$ used in the calculation of $\langle P_T \rangle(0,0)$. The distribution of $\langle P_T \rangle(0,0)$ becomes narrower with increasing sliding window length, though it is unclear whether that is due to fewer samples or a stabilization of the variability on long enough timescales, or both.
We can use these results to extrapolate the PSD measurements from finite to infinite $\tau$. For these accretion flow parameters, the measurement of the distribution of $\langle P_{\tau}\rangle(0, 0)$ calculated using multiple realizations of total length only $\tau \gtrsim 1500 \text{ GM/c}^2$ contains the infinite $\tau$ limit within the distribution of these realizations. For $\tau \lesssim 1500 \text{ GM/c}$, this empirical extrapolation is only valid up to some maximum value of $T$.

Assuming that the conclusions found for $\langle P_{\tau}\rangle(0, 0)$ hold for $\langle P_{\tau\theta}\rangle(u, v)$ and $\langle P_{\tau\phi}\rangle(u, v)$, we now characterize how the uncertainty associated with a single turbulent realization behaves across different baseline lengths. The left column of Figure 17 shows the same procedure as Figure 16 for nonzero baselines. A slight difference exists, since in this test we do not require a constant simulation cadence, and it includes images from earlier in the simulation. No clear trend exists with starting time, indicating that the variability is not secularly evolving. These PSDs combined form a distribution, whose width we characterize as a scatter by standard deviation of the logarithm of the variability PSDs. The right column of Figure 17 shows the same procedure as Figure 16 for nonzero baselines. The median value over these simulations is shown along $v = 0$ (blue) and $u = 0$ (red), and shaded bands show the entire range over 81 simulations.

Figure 17. Effect of simulation length on the variability PSDs. The left column shows $\langle P_{\tau}\rangle(0, v)$ (top) and $\langle P_{\tau\phi}\rangle(0, v)$ (bottom) for a MAD simulation from model set C with $u_{\infty} = 0.9$, $R_{\text{high}} = 20$, and $i = 10^\circ$ with $T = 30,000 \text{ GM/c}$. This simulation has been split up into 18 windows of length 5000 GM/c$^2$. No trend exists with window start time, but a scatter exists. The right column shows a measurement of this scatter for every simulation in model set C along $v = 0$ (blue) and $u = 0$ (red). The median value over these simulations is shown along $v = 0$ (cyan) and $u = 0$ (orange), and shaded bands show the entire range over 81 simulations.

$R_{\text{high}} = 20$, and other accretion flow parameters could show different distribution widths.

A.4. Code Comparison

GRMHD simulations include many numerical approximations and choices that vary between codes. To test the effects of these different implementations, we compare the variability PSDs from model sets A (using only inclinations less than $90^\circ$) and D. These come from different fluid modeling codes (KHARMA and BHAC) and different GRRT codes (ipole and BHOSS) and contain 200 simulations with matching fluid and imaging parameters (see Section 2.2). Each simulation has three windows in model set A and two windows in model set D. To compare these, we introduce a difference measure,

$$D(x, y) = 2 \left( \left| \log_{10} \frac{x}{y} \right| \right)^\text{windows},$$

(A1)

where $x$ and $y$ stand in for one of the variability PSDs for two simulations, and the averaging is performed over all the possible window combinations. This measure is a natural extension of the standard deviation in the log domain in the limit of two data points and so is analogous to that used in Figure 17. Differences between codes could also stem from simulation segments being separate turbulence realizations, so we can compare the difference between windows in separate codes to that between windows in one code. When comparing within the same code, the averaging only includes independent window combinations.

This difference measure is shown for all simulations in Figure 18. The red and green regions show the range of differences between separate windows in the same simulations, which arises...
from separate windows corresponding to different turbulence realizations. These are similar to the regions in Figure 17 and include a larger sampling of GRMHD parameters, but with only two or three measurements compared per simulation.

The blue regions show the range of difference between the two model sets. On short baselines \(|u| \lesssim 10\) G\(\lambda\), the intercode difference and intracode difference are similar, with the median intercode difference somewhat larger. On these large scales, the choices in each code are subdominant to the uncertainty in the PSD measurements from a single realization. On longer baselines, the intercode differences are larger. For \(\langle \hat{P}_T \rangle\), the typical (median) deviation between codes is 1/10 of an order of magnitude (i.e., 25\%).

A comparison of the difference in \(\langle \hat{P}_T \rangle\) on short baselines among these model sets can also be seen through the difference of the centroid of emission in the images in Appendix C.

The intercode differences stem from a variety of numerical and algorithmic choices, some of which have been studied and quantified for GRMHD in Porth et al. (2019) and GRRT in Gold et al. (2020), and we leave further study of code differences for future works.

### Appendix B

#### Short-baseline PSD and Observational Systematic Uncertainties

The short-baseline expansion made in Equation (17) faces a number of practical complications in standard VLBI analysis applications. These include the following:

1. The time-averaged truth image is not known a priori, and thus, typically, \(\tilde{I}_T\) is unavailable. As a result, any approximation of the mean is not strictly constrained in the same manner as \(\tilde{I}_T\) to \(I\).

2. Each visibility is subject to individual statistical uncertainties (e.g., thermal errors) that impose a floor on the variability associated with the measurement details.

3. In the absence of an absolute phase calibration, unknown atmospheric delays add a large (many \(2\pi\)) random station-based component to the phases of the measured visibilities. These phase shifts vary on the atmospheric coherence time, which at 230 GHz can be as short as tens of seconds.

The last of these presents a significant complication for the interpretation of observational estimates of the variability PSDs.

In the absence of a separate phase reference, for VLBI arrays the individual station phase delays are typically set via self-calibration: an optimization process in which the individual station phases are estimated during image reconstruction that leverages the image priors (e.g., positivity) and the fact that the number of baselines, \(N(N-1)/2\), is larger than the number of unknown station phase delays, \(N\), when \(N > 3\), rendering the problem overconstrained. Nevertheless, this procedure is formally degenerate with an overall phase shift and linear phase gradient. The former is benign, and we will not consider it further. The latter, a linear phase gradient, is equivalent to translating the source in the image domain. It is for this reason that, in the absence of a phase reference, VLBI does not have the capacity to constrain the absolute position of the source being imaged.

Additional practical limitations include additional sources of variability (e.g., interstellar scattering), station gain amplitudes, and polarization leakage, all of which will result in effective mismatches between the true intrinsic \(V\) and that observed. We will, however, neglect these in favor of addressing the phase shifts.

The above considerations enter into the formalism of Section 3.2 in two ways: first, the conceptual seving of \(\tilde{I}_T\) and \(I\), and second, the introduction of additional time-dependent phase shifts of \(V\). That is, the observed visibilities are

\[
V_{\text{obs}} = V e^{2\pi i u \cdot x} \tag{B1}
\]

for some arbitrary \(X(t)\) that describe the unknown phase shifts. As a result, the observed intensity map is shifted by \(X\), i.e., \(I_{\text{obs}}(x) = \tilde{I}(x + X)\).

We will require the first four terms (but only the first three explicitly) in the Fourier series expansion employed in Equation (17), i.e., after expanding

\[
e^{-2\pi i u \cdot x} = 1 - 2\pi i u \cdot x - 2\pi^2 (u \cdot x)^2 + \frac{4\pi^3}{3} (u \cdot x)^3 + \ldots \tag{B2}
\]

with associated contributions to the complex visibility:

\[
\begin{align*}
V_0 &= \int (\tilde{I}(x + X) - \tilde{I}_T(x))d^2x = 0 \\
V_1 &= -\int 2\pi i (u \cdot x)(\tilde{I}(x + X) - \tilde{I}_T(x))d^2x \\
&= -2\pi i u \cdot (X - \tilde{X}_T) \\
V_2 &= -\int 2\pi^2 (u \cdot x)^2[\tilde{I}(x + X) - \tilde{I}_T(x)]d^2x \\
&= -2\pi^2 u^T C u + \frac{\nu^2}{2} \\
V_3 &= \int 4\pi^3 (u \cdot x)^3[\tilde{I}(x + X) - \tilde{I}_T(x)]d^2x, \tag{B3}
\end{align*}
\]

where \(X\) and \(\tilde{X}_T\) have their meanings from Section 3.2 and \(C\) is the second moment of \(\tilde{I}(x + X) - \tilde{I}_T(x)\). Therefore, to order \(|u|^4\),

\[
\langle \hat{P}_T \rangle = \langle |V_1|^2 + |V_2|^2 + V_3^* V_1 + V_3^* V_1 \rangle \\
= (4\pi^2 |u \cdot (X - \tilde{X}_T)|^2 + 4\pi^4 (u^T C u)^2) \\
+ Q u \cdot (X - \tilde{X}_T), \tag{B4}
\]

where \(Q\) is a collection of nonsingular terms that are of order \(|u|^3\).

In principle, in the absence of additional information, the value of \(X\) is arbitrary. In practice, \(X\) will be selected on the basis of the prior assumptions underlying the analysis, e.g., temporal regularization or minimization of the deviations from a static model. When they are chosen such that \(X = X - \tilde{X}_T\), the quadratic terms are eliminated altogether, leaving the short-baseline behavior of the average residual PSD completely determined by the quartic term, \(\langle \hat{P}_T \rangle = 4\pi^4 (u^T C u)^2\). Where \(X = X - \tilde{X}_T\), the short-baseline PSD is reflective of the manner in which this is selected.

### Appendix C

#### GRMHD Simulation Images and Their Centroids

Figures 19, 20, and 21 provide average images of the GRMHD simulations in model set A corresponding to the last window and only showing inclinations of 10°, 50°, and 90°. Overlaid on the average images are the 1σ covariance ellipses of the centroid of emission for model sets A and D, scaled larger by a factor of 10. These directly related to the trends discussed in Section 5.2.
**Figure 19.** Average images of the GRMHD simulations at 10° inclination. Ellipses show 1σ contours of the distribution of the centroid of emission for each window in model sets A (solid green) and D (dotted cyan), scaled up by a factor of 10 for visual clarity. Average images are for the last window of model set A only. The black hole axis is pointed up for positive spin and down for negative spin. Color corresponds to intensity per pixel and is not the same among images.
Figure 20. Similar to Figure 19, but for 50° inclination.
Figure 21. Similar to Figure 19, but for 90° inclination.