Constraining Fundamental Physics with the Event Horizon Telescope

Markus Rummel[a,b] **and C.P. Burgess**[a,b]

[a]Physics & Astronomy, McMaster University, Hamilton, ON, Canada, L8S 4M1
[b]Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada

E-mail: rummelm@mcmaster.ca, cburgess@perimeterinstitute.ca

Abstract. We show how Event Horizon Telescope (EHT) observations of the supermassive object at the center of M87 can constrain deviations from General Relativity (GR) in a relatively model-independent way. We focus on the class of theories whose deviations from GR modify black holes into alternative compact objects whose properties approach those of an ordinary black hole sufficiently far from the would-be event horizon. We examine this class for two reasons: (i) they tend to reproduce black-hole expectations for astrophysical accretion disks (and so do not undermine the evidence linking black holes to active galactic nuclei); (ii) they lend themselves to a robust effective-field-theory treatment that expands in powers of ℓ/r, where ℓ is the fundamental length scale that sets the distance over which deviations from GR are significant and r is a measure of distance from the would-be horizon. At leading order the observational impact of these types of theories arise as modifications to the transmission and reflection coefficients of modes as they approach the horizon. We show how EHT observations can constrain this reflection coefficient, assuming only that the deviations from GR are small enough to be treated perturbatively. Our preliminary analysis indicates that such reflection coefficients can already be constrained to be less than of order 10%, and so can rule out some benchmark cases used when seeking black-hole echoes. The precise bounds depend on the black hole spin, as well as on detailed properties of the reflection coefficient (such as its dependence on angular direction).

Contents

1 Introduction

New observational information very often drives progress in fundamental science, and the strikingly new information of our time comes from the recent imaging of black holes using both gravitational waves [1] and the Event Horizon Telescope (EHT) [2]. In principle, the apparent agreement between these observations and the predictions of General Relativity (GR) provide constraints on the way new fundamental physics might deviate from GR in the strong-gravity near-horizon regime, though the reliable extraction of these constraints remains a relatively new field.

Throughout most of physics new observations are typically used to constrain fundamental physics in one of two ways.

- *Model Building:* The first approach considers each detailed proposal for new physics on its own terms, computing its implications on a model-by-model basis. This is the approach used, for instance, when the implications of specific proposals (*e.g.* Brans-Dicke scalars, axion models, *etc.*) are compared with the predictions of GR.

- *Effective Field Theories (EFTs):* The second approach is appropriate if no new degrees of freedom in the modification actually appear in the phenomenon of interest (perhaps they are too heavy or short-ranged to be relevant to the scales being measured). In this case all new degrees of freedom can be integrated out, with new effects parameterized using a low-energy/long-distance expansion. In this framework new physics enters only

through short-distance effects (smaller than some characteristic scale ℓ) and builds local interactions using only known fields, organized in powers of derivatives times ℓ.

A strength of the model-building approach (which is the direction to this point most explored in the literature (see, for example, [3, 4] for surveys with references) is that individual models can be very predictive, with many possible observable deviations from GR characterized by a small number of model parameters. Its main drawback is inefficiency; there are a great many models from which to choose and at present we have no idea which is likely to be the right one. So in principle one must work through them all, making detailed predictions for each. In practice this means a few popular models get detailed attention and others – though perhaps equally deserving – do not.

The strengths (and weaknesses) of an EFT approach are complementary to those of model-by-model calculations. The main strength is the robustness of results: EFTs capture the dominant way that *any* model in a very broad category can affect observations (for reviews relevant to applications with gravity see [5–8]). The broad category of theories whose low-energy effects are described by EFTs must only: (*a*) satisfy the assumption that it not introduce new degrees of freedom at the length scales (call them r) relevant for the observations; and (*b*) have a significant hierarchy, $\ell \ll r$, between r and the new-physics scales (collectively represented by ℓ). The main drawback of EFTs relative to a model-by-model approach is a comparative lack of predictability; an EFT's effective interactions are all *a-priori* independent of one another, while the predictions of any particular model would express them in terms of the smaller set of model parameters.

Because these two approaches have complementary strengths it is usually the most informative when both are deployed. It is the comparison between specific model predictions and generic EFT expectations that most quickly focuses attention on what the most promising directions consistent with observations might be.

Black holes *v.s.* compact sources: a new type of EFT

Although EFTs can play an important role in the post-Newtonian regime of black-hole mergers [9], comparatively little work as been done computing the predictions of EFTs in the strong-gravity, near-horizon regime for black holes. There are two main reasons for this.

The first reason is a practical one: EFTs typically involve a local expansion in powers of derivatives (both curvatures and derivatives of any other fields), like

$$S_{\text{eff}} = - \int \mathrm{d}^4 x \sqrt{-g} \left[-\frac{M_p^2}{2} \mathcal{R} + (\partial \phi)^2 + \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + c_1 \mathcal{R}^2 + c_2 \mathcal{R} F_{\mu\nu} F^{\mu\nu} + c_3 (\partial \phi)^4 + \cdots \right], \tag{1.1}$$

where $\mathcal{R}$ is the Ricci scalar, $F_{\mu\nu}$ is a gauge field strength (such as for electromagnetism) and ϕ is any scalar fields that might be entertained at the energies of interest. The ellipses here involve all possible powers of curvatures and fields and their derivatives. These actions inevitably involve the presence of higher-derivative interactions. Techniques are only now being developed [10–13] to handle efficiently such interactions in the strong-gravity, near-horizon regime. The problem is that higher derivative interactions introduce spurious solutions that have nothing to do with the full theory's low-energy limit. These are known not to cause problems of principle for EFTs (since the spurious solutions do not arise at fixed orders in the expansion in powers of ℓ/r — see *e.g.* the discussions in [8, 14, 15]) this is only cold comfort for numerical calculations, where it is difficult in practice to separate the spurious evolution from real predictions of the low-energy regime.

The second obstruction to using EFT methods in the strong-gravity, near-horizon regime arises because many ideas about small-distance extensions of GR are quite speculative. Although often well-motivated, the resulting theoretical frameworks are usually insufficiently developed to make predictions that are precise enough for comparison with observations. For instance, in the strong-gravity regime relevant to black holes, some such theories propose qualitative changes like the appearance of an enormous number of new degrees of freedom as one approaches the black hole horizon [16–19], but without a concrete framework that allows phenomenological testing. It is usually assumed that these theories only deviate from GR very close to the would-be event horizon, in order not to alter the current understanding of the astrophysics of accretion disks.

As pointed out in [20], theories that only modify GR in the near-horizon regime can be systematically confronted with observations using a slightly different EFT approach to modifications of GR. This approach exploits an expansion in powers of ℓ/r where ℓ characterizes the length scale over which modifications are significant and r is a measure of the distance scale of interest for the observations. (For the applications of this paper the distances of interest prove to be of order the horizon size, $r \sim r_H$.) It is a special case of a more general framework [21–24] that captures how boundaries or compact objects affect their larger surroundings. For such theories the implications of modifications can be explored using an effective action that is localized near the horizon, expanded in powers of ordinary fields and their derivatives. Because EFTs capture the the low-energy implications of *any* possible ultraviolet (UV) extension, confronting the EFT with observations allows the implications of these theories to be explored even though a detailed UV completion is not yet known.

In this paper, we extend the application of these techniques beyond the applications to LIGO considered in [20] to include recent results from the EHT [2]. EHT observations of the event horizon are consistent with M87 being a supermassive black hole described by the Kerr metric, *i.e.* a black hole that is purely described by its mass and spin as in GR. We argue that this agreement can be used to constrain theories of modified gravity and in particular those whose effects are localized near the event horizon.

Reflection, transmission and experiments

As shown in [20], at lowest nontrivial order in ℓ/r the observational implications of theories with exclusively near-horizon deviations from GR can be described in terms of a near-horizon reflection coefficient, R, where $|R|^2$ represents the reflection probability for an inward-directed wave[1] 'at' the horizon. GR emerges as the limit $R \to 0$, where every wave is purely infalling sufficiently near the horizon. For more general R a nonzero fraction of an incoming wave is reflected at the would-be event horizon, and it is the implications of these that provide potential new-physics signatures.[2] The relevance of near-horizon reflection to such theories is in any case intuitive, and the hypothesis of nonzero R was earlier used[3] to identify interesting

---

[1]More precisely, partially reflecting boundary conditions are imposed, say, on a surface just outside the horizon, such as at radial position $r = r_H + \epsilon$. The discussion below reviews the arguments of [20] as to why nothing physical depend on the precise choice of ϵ.

[2]The replacement of UV physics with boundary conditions also applies elsewhere in physics, such as the influence of nuclear structure on atomic energy levels (where it is the small ratio, r_N/a_B, of nuclear to atomic size that controls the EFT). The observation that nuclear properties affect atomic energy levels only through such boundary conditions allows them to be computed fairly efficiently [25].

[3]Earlier work tends not to distinguish the scales ϵ and ℓ, which in general can be very different. For applications to the effects of nuclear structure on atomic energy levels the RG-invariant scale is $\ell \sim (Z\alpha)^2 r_N$; much smaller than nuclear size, r_N, despite choosing boundary conditions outside the nucleus: $\epsilon > r_N$ [23, 25].

signatures such as for gravitational-wave echoes [26–30].

In the context of the EHT, the deviations from GR that are captured by a reflection coefficient can be constrained in the following way: the EHT observes radio waves that are emitted as synchrotron radiation from the relativistic electrons in the hot magnetized optically thin accretion disk of M87. If there is a sufficient amount of reflection near the horizon the image as seen by the EHT deviates from the predictions of GR: the radio waves emitted by the accretion disk that fall into the black hole in vanilla GR become instead partly reflected and escape. Some of these waves would then survive to be seen by the EHT. In principle, sufficiently large deviations between the observed EHT picture and what is consistent with a vanilla black hole can be excluded. This type of logic has been used to constrain several theories of modified gravity, such as naked singularities [31, 32], extra dimensions [33, 34] and scalar hair [35, 36]. The focus of this work is to follow [20] and show how to constrain *all* UV modifications of GR whose implications are localized near the horizon, and to use the EFT framework to quantify the theoretical error that is involved in making potentially-observable predictions.

As a proof of concept, we implement the program of constraining different reflection coefficients in a particularly simple way. As is standard in these calculations, we imagine a near-horizon reflection coefficient to have been specified on a surface just outside the horizon.[4] (Although much of the literature specializes to Schwarzschild geometries, we here consider a reflection coefficient — in both spherically symmetric and direction-dependent versions — specified on a surface just outside a Kerr black-hole horizon.[5]) Although some earlier workers perform their phenomenological analyses using relatively extreme reflection probabilities, like $|R| = 1$ — *i.e.* perfect reflection — we instead keep this parameter general since this allows us to quantify how big R can be and remain consistent with observations. We show below that (for electromagnetic waves, at least) extreme choices like $|R| = 1$ are very likely already ruled out by the EHT.

Modelling the compact object's environment

In principle, to constrain the existence of (or discover) new physics one must simply compute the image expected with GR and the image expected in the presence of near-horizon reflection, and ask whether these can be distinguished by EHT observations. In practice, the great complication comes from modelling the compact object's environment, since this is ultimately the source of the light being imaged. For M87 this environment is believed to be an optically thin accretion disk [37–39], whose detailed modelling is required when constructing and interpreting the EHT images [40–42].

Rather than providing a similarly detailed model of this environment, we here take a simpler approach that builds on the fact that the observations seem in first approximation to be well-described by GR. We obtain our constraint by starting with an EHT image and tracing the observed light rays back to the vicinity of M87 under the assumption that they pass through the vanilla Kerr black-hole geometry predicted by GR. We ignore light scattering when doing so because the black-hole environment is optically thin. This ray-tracing exercise allows us to map the observed light intensity back to a light-intensity map on any particular reference surface that surrounds the compact object. To see the effects of reflection we then re-

---

[4]In practical examples we choose this to be a surface of fixed coordinate radius at $r = r_H + \epsilon$.

[5]One might worry that physical results depend on the precise position of this surface, but it turns out the precise position of this surface is irrelevant because physical observables do not depend on it. This does not however preclude them from depending on the physical (but distinct) length-scale ℓ [20].

propagate these light rays back out to an observer at infinity,[6] again using the Kerr geometry, but this time assuming a nonzero near-horizon reflection coefficient, R, on a surface very near to the position of the black-hole horizon. Sufficiently much reflection of this type alters the intensity map seen at infinity (because, for example, light that otherwise might have fallen into the black hole now gets reflected and escapes to infinity), causing it to differ from the starting EHT image. The picture of the reflected rays is then added to the original EHT picture and if the combined picture is inconsistent with the diagnostics established by the EHT collaboration [40–42], this particular reflection coefficient can be excluded.

There are three key assumptions in this reasoning. The first is that the medium is optically thin, so that ray intensity does not vary appreciably while en route to the observer. The second is that any particular element of the reference surface radiates equal intensity into all directions, so that the intensity of the reflected rays depends only on where on the reference surface it starts (together with the precise value of the reflection coefficient). The third assumption is that the effects of reflection are small; there is not enough of it to feed back on the environment and alter the intensity distribution inferred in the absence of reflection. Our explicit ray-tracing code is available online at https://github.com/mrummphys/EventHorizonTelescope.

We find that we can typically constrain reflection coefficients to be less than of order $1 - 10\%$. These constraints only weakly depend on the so-far poorly constrained spin of the compact object and the angular dependence assumed for R. The most constraining image statistics turn out to be the deviation from circularity of the EFT image's ring-like structure and, to a lesser extent, its fractional central brightness (*i.e.* the ratio of intensity in the center of the image to the intensity in the ring-like structure).

We emphasize that our analysis presented here is based only on a digitized image of the EHT observations. One should be able to improve these bounds considerably using the original (and future) EFT dataset, and by incorporating reflection directly into the relativistic magnetohydrodynamic (GRMHD) simulations used by [41] to model the compact object's environment. Although these extra steps can and should be done, our analysis both provides a demonstration of concept and a relatively simple way to estimate the size of EHT constraints on classes of deviations from GR. We hope to address some of these issues in future work.

A road map

This paper is structured as follows: in Section 2, we introduce reflection coefficients as described by the EFT we use for near-horizon physics. In Section 3, we introduce the EHT observations and image analysis techniques used by the EHT collaboration followed by a description of our ray tracing and image creation ethology in Section 4. The results are discussed in Section 5, followed by our conclusions in Section 6. Appendix A recaps the connection between R and the fundamental-physics scale ℓ.

2   The near-horizon EFT analysis

An interesting class of approaches to black-hole information-loss problems argue that deviations from GR arise only as an observer approaches the event horizon [16–18]. For example, one proposal in this category argues black holes are not really black; instead they just involve a very large number of new microscopic degrees of freedom whose huge degeneracy is related

---

[6]Strictly speaking, we actually re-propagate rays back in to the compact object from infinity to find where on the reference surface they arrive.

to the enormous black hole entropy. Although it is usually hard to make precise predictions in these proposals about the properties of astrophysical black holes, these new degrees of freedom are expected to become accessible only close to where the would-be black hole horizon forms [19, 43].

Having any hypothetical new physics be localized to lie within a distance ℓ of the horizon, with ℓ much smaller than the horizon size itself, $\ell \ll r_H$, would be in any case also attractive since such a construction would have been relatively easy to miss until the present. Both of these reasons argue for quantifying the observational implications of such theories for the new observational windows into the near-horizon regime. A low-energy EFT approach would be particularly valuable since its validity can be much broader than any particular (as yet, possibly, ill-defined) UV completion. Because they involve systematic low-energy expansions, EFT methods also lend themselves to explicitly quantifying any theoretical uncertainties intrinsic to the predictions. We now sketch the construction of the EFT for exotic near-horizon physics, and reproduce the argument of ref. [20] that its leading effect is to give a reflection coefficient (and, sometimes, a damping time) for any particular mode.

2.1 EFTs and near-horizon boundary conditions

The starting point is the effective action. When ℓ is much smaller than the scales of practical interest (such as the horizon size r_H) then it is useful to integrate out all of the new physics associated with any new degrees of freedom localized near Σ. Once this is done the influence of this new physics is captured by a component of the effective action that is localized on a surface very near the horizon. For example, for a single complex Klein-Gordon scalar field in a black-hole background this leads to an action of the form $S = S_B + S_{\rm hor}$ where

$$S_B = -\int {\rm d}^4 x \sqrt{-g}\Big[g^{\mu\nu}\partial_\mu\phi^* \, \partial_\nu\phi + m^2\phi^*\phi + \cdots\Big] \tag{2.1}$$

and

$$S_{\rm hor} = -\int_\Sigma {\rm d}^3 x \sqrt{-\gamma}\left[h_0 + h_1\,\phi^*\phi + \cdots\right], \tag{2.2}$$

where the integration in (2.2) is over a time-like surface, Σ, defined by $x^\mu = y^\mu(\sigma^a)$ just outside the horizon (e.g. perhaps $r = r_H + \epsilon$ for a Schwarzschild black hole), and $\gamma_{ab} = g_{\mu\nu}\partial_a y^\mu \, \partial_b y^\nu$ is the induced metric on Σ.

For both S_B and $S_{\rm hor}$ the ellipses denote all possible local combinations involving more powers of both the field and its derivatives. The more fields or derivatives appearing in such terms, the more suppressed by powers of the microscopic length scale ℓ their effective couplings must be. For instance, for a canonically normalized scalar like ϕ the coupling h_1 in (2.2) has dimensions (length)$^{-1}$. An effective interaction in S_b proportional to $h_n(\phi^*\phi)^n$, on the other hand, would have a coupling h_n with dimension (length)$^{2n-3}$. The reality of the effective coupling h_n is related to probability conservation. If overall probability is conserved at the horizon, h_n is real but if not – as in the black hole case – h_n is complex with the imaginary part of h_n quantifying probability loss/absorption of the system [20].

Effective couplings (like h_1) appearing in S_b affect the dynamics of ϕ away from Σ only through the boundary condition that they contribute near Σ. This can be seen most easily in the semiclassical limit when examining the saddle point for functional integrations over ϕ (both on and off the surface Σ). Variations off the surface (in the 'bulk') reproduce the classical field equation, $(-\Box + m^2)\phi = 0$, while demanding a saddle point for $S_B + S_{\rm hor}$ right

at Σ give the boundary condition

$$\left(n_\mu g^{\mu\nu}\partial_\nu\phi + h_1\phi + \cdots\right)_\Sigma = 0\,, \tag{2.3}$$

where n_μ is the unit outward-pointing normal to Σ and ellipses denote the influence of any terms hidden in the ellipses of eq. (2.2).

This boundary condition (2.3) can be interpreted in two complementary ways. First, the value of h_1 dictates the normal derivative of $\ln\phi$ on Σ, which in turn dictates one of the integration constants found when integrating the Klein Gordon equation in the bulk. Since these integration constants in turn control the relative size of ingoing and outgoing modes, a straightforward argument shows that knowledge of the pair (h_1, ϵ) is equivalent to knowledge of the reflection coefficient $R[h_1, \epsilon]$. Although this makes it seem as if R depends on ϵ, this is not really true because the value of h_1 found by integrating out UV physics is itself ϵ-dependent, in just such a way that $(\mathrm{d}/\mathrm{d}\epsilon)R[h_1(\epsilon), \epsilon] = 0$. That is, the functional form of $h_1(\epsilon)$ is precisely what is required to ensure that R does not depend on ϵ.

This is a special case of a general EFT argument: the regularization scale ϵ drops out of observables because it is absorbed into a renormalization of $h_1(\epsilon)$, defining an RG flow $h_1(\epsilon/\ell)$, for some RG-invariant scale ℓ. This provides the second, complementary, way to read eq. (2.3): because it holds for *any* value of ϵ it remains true even after differentiation with respect to ϵ. If this is done with the reflection coefficient held fixed it provides the precise form for $h_1(\epsilon)$ required to ensure that R is ϵ-independent. This allows $R[h_1(\epsilon/\ell), \epsilon]$ to be traded for a more informative relationship $R = R(\ell/r_H)$ between R and the physical RG-invariant length-scale ℓ (given in more detail in Appendix A.) Upper limits to $|R|$ in this way turn into upper limits on ℓ/r_H.

For black holes the infrared fixed point of the RG flow one finds in this way for h_1 is complex and corresponds to the choice of purely infalling boundary conditions at the horizon, *i.e.* to vanilla GR [20]. Details of this construction are given in [20–23] and so are not repeated here. Instead, for the purposes of EHT phenomenology we simply imagine the reflection coefficient R to have been specified.

2.2 Reflection Coefficients

Given that a class of modifications of GR can be described by replacing the perfect-infall boundary condition of GR with a mixture of incoming and outgoing wave near the horizon, we next turn to what potentially observable consequences this might have. As has been pointed out, partial reflection of incoming gravitational waves near the horizon can lead to a series of echoes that might be observable at LIGO [26–29, 44]. In this paper, we instead explore the implications of this kind of boundary condition for electromagnetic waves as probed by the EHT.

To study reflection we need to specify a reflection coefficient R at a surface Σ, which for simplicity we choose to be a surface of fixed coordinate radius, $r = r_R = r_H + \epsilon$. As discussed above, the value of r_R is not physically relevant as R is RG invariant. In specifying R we select a particular representation of an RG flow representing a particular UV completion. We will look into three different kind of possibly angular dependent reflection coefficients inspired by a multipole expansion:

$$\begin{aligned} l = 0: \quad & R(\theta, \phi) = R_0\,, \\ l = 1: \quad & R(\theta, \phi) = R_0\,|\cos(\theta)|\,, \\ l = 2: \quad & R(\theta, \phi) = R_0\,|\sin(\theta)\cos(\theta)|\,, \end{aligned} \tag{2.4}$$

where the higher multipoles $l = 1, 2$ introduce a dependence on the polar angle θ. One might expect more complicated angular dependence in a particular UV modification but (2.4) is a starting point for the simplest reflection coefficients one might expect. For a discussion how the angular dependence of R is related to the angular dependence of the point-particle-effective-field-theory (PPEFT) coupling describing the black hole, see [20] and Appendix A. Finally, R might also be frequency dependent in a generic UV modification, i.e. a waves' incoming frequency might be different from the reflected outgoing waves' frequency due to energy absorption/emission by the black hole object. We do not take this scenario into account in this work and instead focus on the simpler scenario described by (2.4).

3 EHT observations

The EHT [2] is a very long baseline interferometry experiment that measures radio brightness distributions at a wavelength of 1.3 mm on the sky. Its unprecedented angular resolution allows to resolve structures of angular scales of $\mathcal{O}(\mu\text{as})$. This makes it possible for the first time to resolve event-horizon-scale physics of supermassive black holes that are relatively nearby and/or active, i.e. bright. In particular, the EHT collaboration was able to reveal the shadow caused by gravitational light bending and photon capture at the event horizon of the supermassive black hole at the center of the giant elliptical galaxy M87 at a distance of 16.8 Mpc [45]. The images of M87 show a ring-like structure with a diameter of about 40 μas with a central brightness depression due to the event horizon of the black hole. The ring brightness is asymmetric which can be explained by relativistic beaming of the photons emitted as synchrotron radiation from the plasma rotating in the accretion disk at close to the speed of light. The angular size of the ring is directly related to the mass (and to a lesser extend spin) of the supermassive black hole and is estimated by the EHT collaboration as $M = (6.5 \pm 0.7) \cdot 10^9 M_\odot$ [2]. There is a slight tension with estimates of the mass via stellar dynamics at $M = 3.5^{+0.9}_{-0.3} \cdot 10^9 M_\odot$ [46].

The images of the EHT collaboration are consistent with a shadow of a rotating Kerr black hole in general relativity. The EHT observations offer a new unique opportunity to test near horizon gravitational physics, such as extensions to general relativity described in Section 2.

In order to constrain modified theories of gravity we first need an image of the EHT observations that we can compare the modified gravity images to. Since the EHT observations are consistent with general relativity, i.e. a Kerr black hole, a modified theory of gravity can be excluded if its image is too dissimilar from the EHT/Kerr image. To determine if two images are dissimilar we invoke the image analysis variables that we will introduce in Section 3.1.

As we do not have access to the EHT data, we digitize one of the EHT images. Since this is a proof-of-concept study a digitized version will be good enough for our purposes however we generally expect our bounds derived in Section 5 to be more sensitive by a factor of a few with the original data. A generic representative of the different imaging methods `DIFMAP`, `eht-imaging` and `SMILI` (see [41]) of the EHT is Figure 15 in [41] which is an average of these three methods observed on April 11 2017. To digitize the image, we use the colour code provided in the figure to find all pixels with brightness temperature T in a linear sample between 0 and $6 \cdot 10^9$ K with $5 \cdot 10^8$ K step size. Next, we create a pixel grid of size -55 to 55 μas with pixel size of 1 μas. Each pixel is assigned an intensity via a 2D interpolation function from the above points of the original image. The obtained image is shown in Figure 1.
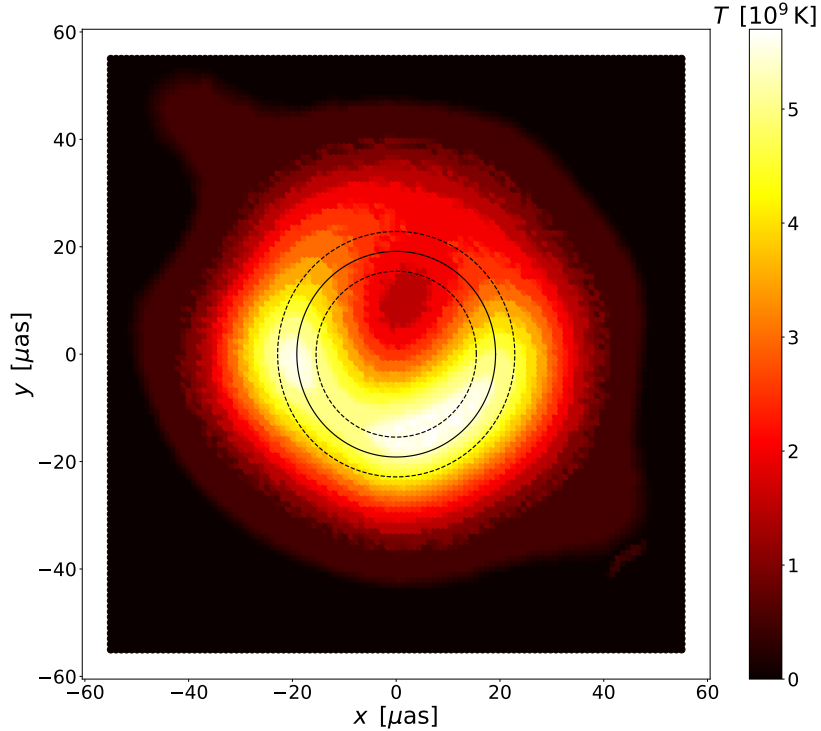
Figure 1. The digitized EHT image. The solid black line is half a ring diameter $d/2$ from the image center where d is defined in eq. (3.3). The dashed black lines are at radii $(d - \sigma_d)/2$ and $(d + \sigma_d)/2$, respectively where σ_d is defined in (3.4).

3.1 Image analysis

Here, we give a brief review of the image diagnostics the EHT collaboration uses to characterize the ring like structure in an image. By comparing these diagnostics of an image from modified gravity to the observed EHT image we can determine if a theory is either excluded or consistent with the experiment.

Following Section 9 in [41], we first have to find the center of the ring like structure which is the position (x_0, y_0) in the image that minimizes the normalized peak dispersion, i.e.

$$(x_0, y_0) = \text{argmin} \left[\frac{\sigma_{\bar{r}}(x, y)}{\bar{r}_{\text{pk}}(x, y)} \right]_{(x, y)}, \tag{3.1}$$

where $\bar{r}_{\text{pk}}(x, y)$ and $\sigma_{\bar{r}}(x, y)$ are the mean and standard deviation of the peak brightness

$$r_{\text{pk}}(\theta; x, y) = \text{argmax}_r \left[I(r, \theta; x, y) \right], \tag{3.2}$$

with respect to θ. $I(r, \theta; x, y)$ is the intensity of the image at pixel location defined by (r, θ) coordinates with respect to an origin at (x, y). In practice, θ and r are linearly sampled over a discrete set of values between 0 and $50\,\mu$as and $0°$ and $360°$ respectively while we select x and y values each from 20 equally spaced values between -8 and $8\,\mu$as. In order to avoid spurious detection, [41] limits the peak finding algorithm to pixels that have at least 95% of the image's peak brightness. Since our digitized version of the EHT observation has about three times worse brightness resolution we loosen this constraint to at least 60% of peak brightness instead.

Once the image center is found the measured diameter and its uncertainty are defined as the mean

$$d = 2\bar{r}_{\mathrm{pk}}(x_0, y_0) \,, \tag{3.3}$$

and standard deviation

$$\sigma_d = 2\sigma_{\bar{r}}(x_0, y_0) \,. \tag{3.4}$$

We will suppress the (x_0, y_0) argument in the following for all quantities. The deviation from circularity is defined as

$$\Delta_c = \frac{\sigma_d}{d} \,, \tag{3.5}$$

which will be a key quantity to constrain modified theories of gravity effectively described by a reflection coefficient in Section 5.

The ring width is defined as

$$w = \langle \mathrm{FWHM}\left[\mathrm{I}(\mathrm{r}, \theta) - \mathrm{I}_{\mathrm{floor}}\right]\rangle_\theta \,, \tag{3.6}$$

where FWHM is full width half maximum over a radial profile at given angle θ. $\mathrm{I}_{\mathrm{floor}} = \langle I(r_{\mathrm{max}} = 50\mu\mathrm{as}, \theta)\rangle_\theta$ is subtracted from the intensity profile to avoid introducing bias between different images. While (3.6) defines the mean we can also calculate σ_w as the standard deviation of the set of FWHMs in (3.6).

The ring orientation angle is defined as

$$\eta = \left\langle \arg\left[\int_0^{2\pi} I(\theta)e^{i\theta}\mathrm{d}\theta\right]\right\rangle_{r\in[r_{\mathrm{in}}, r_{\mathrm{out}}]} \,, \tag{3.7}$$

i.e. the mean over radii between $r_{\mathrm{in}} = (d - w)/2$ and $r_{\mathrm{out}} = (d + w)/2$ of the first angular mode of the angular profile $I(\theta)$. σ_η is defined as the standard deviation over the set of radii in (3.7).

The degree of azimuthal asymmetry is defined as the mean

$$A = \left\langle \frac{\int_0^{2\pi} I(\theta)e^{i\theta}\mathrm{d}\theta}{\int_0^{2\pi} I(\theta)\mathrm{d}\theta}\right\rangle_{r\in[r_{\mathrm{in}}, r_{\mathrm{out}}]} \,, \tag{3.8}$$

with σ_A defined as the corresponding standard deviation. A takes values between 0 and 1 and can be interpreted as how evenly the brightness in a ring is distributed over the azimuth angle. Perfect azimuthal symmetry corresponds to $A = 0$ while a delta function concentrating all brightness at one particular angle corresponds to $A = 1$.

The fractional central brightness is defined as

$$f_c = \frac{\langle I(r, \theta)\rangle_{\theta, r\in[0, 5\mu\mathrm{as}]}}{\langle I(d/2, \theta)\rangle_{\theta\in[0, 2\pi]}} \,, \tag{3.9}$$

i.e. the ratio of the mean of the intensity within a disk of radius 5 μas and the mean brightness around the ring.

For our digitized EHT observation on April 11 we find the image characteristics listed in Table 1 compared to those listed in [41]. Note that we use an overlayed/averaged version of DIFMAP, eht-imaging and SMILI and DIFMAP has the biggest blur.[7] Hence, we mainly have to compare to the DIFMAP row in Table 1 as the averaged image is dominated by its blur.

---

[7]As referred to in [40] the blur of an image depends on the beam size it has been restored with. DIFMAP uses the largest beam size of the three image pipelines at 20μas while eht-imaging and SMILI use 17.1μas and 18.6μas, respectively.

	d (μas)	w (μas)	η (°)	A	Δ_c	f_c
Digitized	38.3 ± 7.4	28.4 ± 7.4	205.1 ± 83.7	0.15 ± 0.08	0.19	0.63
DIFMAP	40.7 ± 2.6	29.0 ± 3.0	173.3 ± 4.8	0.23 ± 0.04	0.06	0.5
eht-imaging	41.0 ± 1.4	15.5 ± 1.8	168.0 ± 6.9	0.20 ± 0.02	0.03	0.04
SMILI	42.3 ± 1.6	15.6 ± 2.2	167.6 ± 2.8	0.22 ± 0.03	0.04	$6 \cdot 10^{-6}$

Table 1. Image characteristics of our digitized image compared to the EHT image pipelines for the April 11 observation.

Our image diagnostics in Table 1 are consistent with those of [41] but our errors are larger as is to be expected from a digitized image. For diameter d, width w and azimuthal asymmetry A the uncertainties are a factor 2 - 3 larger than in DIFMAP while for the orientation angle η the error is much larger. We conclude that our digitized image has about three times worse brightness resolution than the EHT image.

For the deviation from circularity and the fractional central brightness the EHT collaboration reports the upper bounds [41]

$$\Delta_c \lesssim 0.1 \qquad \text{and} \qquad f_c \lesssim 0.5 \qquad \text{(EHT)}, \tag{3.10}$$

as a summarizing result from their different images and pipelines. From our digitized image we can only derive slightly weaker upper bounds on these quantities from Table 1:

$$\Delta_c \lesssim 0.2 \qquad \text{and} \qquad f_c \lesssim 0.7 \qquad \text{(digitized)}, \tag{3.11}$$

where we have rounded up to the next full digit from the results in Table 1 to be conservative. Digitizing the overlayed/averaged image gives us a conservative scenario in constraining Δ_c and f_c as these quantities are the least constrained in DIFMAP which dominates the errors in this overlayed image.

We can now formulate precisely what it means for an image generated from a modified gravity theory to be consistent with EHT observations: An image is consistent if the image characteristics d, w, η and A are in agreement with the values in the first row of Table 1 and the upper bounds in (3.11) are not violated. Otherwise it is inconsistent and the theory is excluded by the EHT observations.

4 Methodology

In order to constrain theories of modified gravity described by a reflection coefficient we create images that are the sum of the original EHT image and an image of rays reflected close to the horizon at r_R. Therefore, we first describe our ray tracing algorithm in Section 4.1 and then explain how the images from modified gravity theories are created in Section 4.2.

4.1 Ray tracing

We need capture the light rays that approach the black hole, get reflected at r_R and finally arrive at the camera. In principal, one possibility would be to consider all initial conditions (position and direction) in the vicinity of the black hole and trace their evolution according to the Kerr geodesic equation to see which light rays arrive at the camera and from which direction. As this is practically impossible a much more efficient computational way to tackle this problem is to trace the light ray backwards in time, i.e. *ray tracing* [47–52]. In this

approach, the initial condition for the geodesic equation is the position and angle at which the light ray arrives at the camera, in this case the EHT. Then, the light ray is evolved backwards in time according to the Kerr geodesic equation. At some point, it might hit the surface of reflection at r_R where it is reflected radially, while the intensity of the ray is modified according to (2.4). From there, it evolves away from the black hole to a place infinitely far away from the black hole. In the following, we describe our procedure of solving the Kerr geodesic equation and its initial conditions.

The geodesic equation for the coordinates $x^\mu(\lambda)$ for an affine parameter λ parametrizing the geodesic can be obtained by varying the line element in GR as

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\partial \epsilon}{\partial \dot{x}^\mu} = \frac{\partial \epsilon}{\partial x^\mu} \,, \tag{4.1}$$

where $\dot{x}^\mu \equiv \mathrm{d}x^\mu/\mathrm{d}\lambda$ and

$$\epsilon = -g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = -g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \,. \tag{4.2}$$

The Kerr metric in Boyer-Lindquist coordinates is

$$\begin{aligned}
\mathrm{d}s^2 = {} & - g_{\mu\nu} \, \mathrm{d}x^\mu \mathrm{d}x^\nu \,, \\
= {} & - \left(1 - \frac{r_s r}{\rho^2}\right) \mathrm{d}t^2 - \frac{2 a r_s r \sin^2\theta}{\rho^2} \mathrm{d}t \mathrm{d}\phi + \frac{\rho^2}{\Delta} \mathrm{d}r^2 + \rho^2 \mathrm{d}\theta^2 \\
& + \sin^2\theta \left[(r^2 + a^2) + \frac{r_s r a^2 \sin^2\theta}{\rho^2} \right] \mathrm{d}\phi^2 \,,
\end{aligned} \tag{4.3}$$

where M is the mass of the black hole, a is the spin in units of M. Furthermore (setting Newton's constant to unity)

$$\begin{aligned}
r_s &= 2M \,, \\
\Delta &= r^2 - 2Mr + a^2 \,, \\
\rho^2 &= r^2 + a^2 \cos^2\theta \,.
\end{aligned} \tag{4.4}$$

The horizon is located at

$$r_H = M + \sqrt{M^2 - a^2} \,. \tag{4.5}$$

Eq. (4.1) are four second order differential equations for four variables (t, r, θ, ϕ). By introducing the momentum variables

$$p^\mu \equiv \dot{x}^\mu \,, \tag{4.6}$$

these can be transformed into 8 first order differential equations. Using energy conservation, the time variables t and p^t can be eliminated from the equations, see e.g. [53], and one is left with 6 first order differential equations for the variables $(r, \theta, \phi, p^r, p^\theta, p^\phi)$

$$\dot{x}^i = p^i \qquad \text{and} \qquad \dot{p}^i = F^i(x^j, p^j) \,, \tag{4.7}$$

where the lengthy expressions $F^i(x^j, p^j)$ are listed in Appendix B. We numerically solve (4.7) for the initial conditions discussed below using the `scipy` package `odeint`. To check our code, we reproduce the results of [50] for Kerr black holes.

The initial conditions for the backwards ray tracing are set at the camera. We choose a Cartesian coordinate with the x-direction parametrizing the distance to M87:

$$\boldsymbol{x}_0 = \begin{pmatrix} r_{\text{cam}} \\ (2h-1)\delta \\ (2v-1)\delta \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{p}_0 = \begin{pmatrix} -E \\ 0 \\ 0 \end{pmatrix} \,, \tag{4.8}$$

where r_{cam} is the distance to M87, δ is half the angular size of the image, $h, v \in [0, 1]$ are $n_{\text{pixel}} = 2\delta/\delta_{\text{pixel}}$ equispaced variables each, where δ_{pixel} is the pixel size, indexing the pixel the light ray originates from. E is the energy of the ray/photon. Note that the rays are chosen to be parallel to the x-axis as the distance to the black hole is much larger than its angular size, in fact about 10^{10} larger as δ is of the order of 10th of μas. For the numerics, this allows us to choose r_{cam} as a radius that is far enough from the black hole for the metric to be sufficiently flat as opposed to the huge physical distance to M87. In other words, the ray simply follows a straight line trajectory parallel to the x-axis from this radius to the camera that we do not have to include in the numerical calculations.

For our ray tracing algorithm we choose the following numerical values for the above parameters:

$$r_{\text{cam}} = 100\mu\text{as}\,, \qquad E = 1\,, \qquad \delta = 30\mu\text{as}\,, \qquad \delta_{\text{pixel}} = 1\mu\text{as}\,. \tag{4.9}$$

Note that the choice for E here is somewhat arbitrary as the shape of the trajectory of a ray is independent of its energy. Finally, we have to choose a value for the black hole mass M which we fix by the angular size of the ring in the EHT image of M87. In [42], the EHT collaboration notes that the ring diameter as defined in (3.3) typically arises at an angular size that is 10% larger than the photon ring, i.e. at a radius of 21.1 μas for our digitized image, using the central value for d from Table 1. For a given black hole spin a we can calculate the size of the photon ring and choose M such that it matches this radius. This gives us values between $M = 5.7 \cdot 10^9 M_\odot$ ($a = 0$) and $M = 6.1 \cdot 10^9 M_\odot$ ($|a| = 0.94$), using a distance of 16.8 Mpc to M87.

In order to numerically solve the differential equations (4.7) we need to transfer the initial position and momentum in the camera frame to Boyer-Lindquist coordinates. This is a two step procedure: first, we need to rotate (4.8) by the orientation angle of the black hole which we set to $\theta_{\text{obs}} = 17°$ [54]. The rotation matrix is:

$$\boldsymbol{R}(\theta_{\text{obs}}) = \begin{pmatrix} \sin\theta_{\text{obs}} & 0 & -\cos\theta_{\text{obs}} \\ 0 & 1 & 0 \\ \cos\theta_{\text{obs}} & 0 & \sin\theta_{\text{obs}} \end{pmatrix}. \tag{4.10}$$

Second, we find the values for position and momentum in Boyer-Lindquist coordinates by inverting the relations

$$\boldsymbol{R} \cdot \boldsymbol{x}_0 = \begin{pmatrix} r_{\text{cam}} \sin\theta_{\text{obs}} \\ 0 \\ r_{\text{cam}} \cos\theta_{\text{obs}} \end{pmatrix} = \begin{pmatrix} \sqrt{r_0^2 + a^2} \sin\theta_0 \cos\phi_0 \\ \sqrt{r_0^2 + a^2} \sin\theta_0 \sin\phi_0 \\ r_0 \cos\theta_0 \end{pmatrix}, \tag{4.11}$$

and

$$\boldsymbol{R} \cdot \boldsymbol{p}_0 = \begin{pmatrix} \frac{r_0}{\sqrt{r_0^2+a^2}} p_0^r \sin\theta_0 \cos\phi_0 + \sqrt{r_0^2 + a^2} p_0^\theta \cos\theta_0 \cos\phi_0 - \sqrt{r_0^2 + a^2} p_0^\phi \sin\theta_0 \sin\phi_0 \\ \frac{r_0}{\sqrt{r_0^2+a^2}} p_0^r \sin\theta_0 \sin\phi_0 + \sqrt{r_0^2 + a^2} p_0^\theta \cos\theta_0 \sin\phi_0 + \sqrt{r_0^2 + a^2} p_0^\phi \sin\theta_0 \cos\phi_0 \\ p_0^r \cos\theta_0 - r_0 p_0^\theta \sin\theta_0 \end{pmatrix}, \tag{4.12}$$

for (r_0, θ_0, ϕ_0) and $(p_0^r, p_0^\theta, p_0^\phi)$.
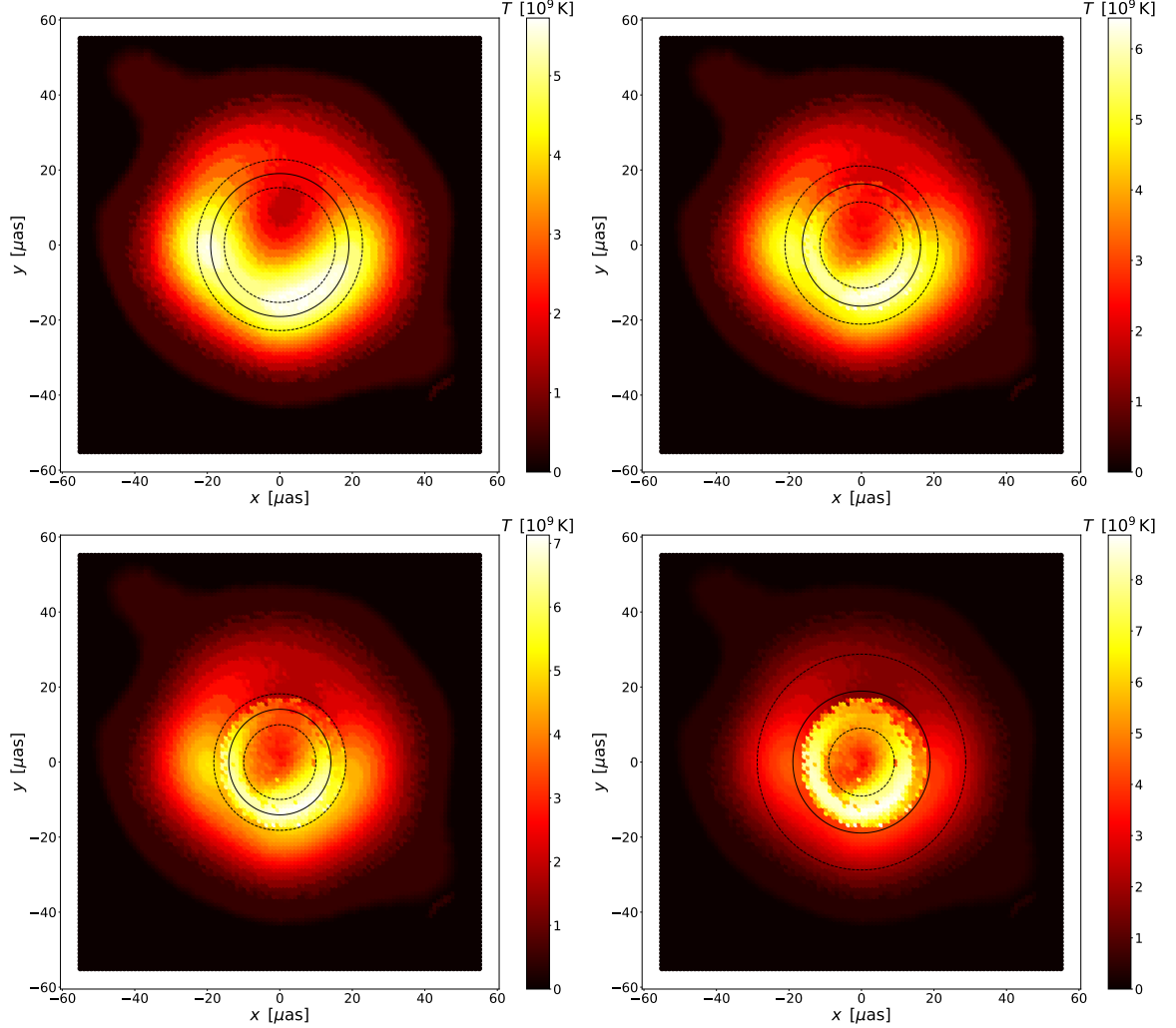
We implement ray reflection at

$$r_R = 1.2\, r_H\,. \tag{4.13}$$

As discussed in Section 2, the choice for r_R is not physical as only R is physical and RG invariant. The reflection condition is imposed as a new initial condition for the differential equation solver if the ray crosses the reflection surface r_R:

$$\left(r, \theta, \phi, p^r, p^\theta, p^\phi\right)_{r=r_R} \rightarrow \left(r, \theta, \phi, -p^r, p^\theta, p^\phi\right)_0 \qquad (4.14)$$

i.e. radial reflection.

4.2 Image creation

Figure 2. The added EHT image for $R_0 = 0.01, 0.2, 0.4, 1$ (top left to bottom right) for $l = 0$ and black hole spin $a = 0.5$. The solid black line is half a ring diameter $d/2$ from the image center where d is defined in eq. (3.3). The dashed black lines are at radii $(d - \sigma_d)/2$ and $(d + \sigma_d)/2$, respectively where σ_d is defined in (3.4).

To create the modified EHT images we implement the following procedure:

- Ray trace the pixels $p \in P$ in the original EHT image back to a surface close to the horizon. This creates a "Close to the Horizon" (CTH) map at radius r_{CTH} that assigns

the intensity of each pixel I_p of the EHT image to a point on this surface $(r_{\mathrm{CTH}}, \theta, \phi)$:

$$\Omega_p \equiv (\theta, \phi)_p = \Omega_p(I_p). \tag{4.15}$$

We choose r_{CTH} to be the photon ring radius, i.e. the radius at which photons travel in unstable orbits. Hence, the horizon, reflection surface and CTH radii obey the hierarchy

$$r_H < r_R < r_{\mathrm{CTH}}. \tag{4.16}$$

The CTH map allows us to assign intensities to rays that have been reflected close to the horizon in a modified theory of gravity as it gives us an idea where the rays we observe at the EHT originate in close proximity to the supermassive black hole.

- We now ray-trace the pixels turning on a reflection coefficient: Starting at the camera within the photon ring a ray would have simply fallen into the horizon in GR. However, with a non-vanishing reflection coefficient the ray gets (partially) reflected where it hits the surface r_R and ultimately escapes to infinity. What intensity do we assign to this particular ray? Somewhere between reflection and the escape to infinity, at Ω_{CTH} the ray crosses the CTH surface. As the CTH map informs us about the origin of a rays intensity we can use this map to assign an intensity of I_R to the reflected ray as the intensity I_p of the closest point to crossing on the CTH surface:

$$I_R = R(\Omega_{\mathrm{CTH}}) \cdot \arg \left[\min \left\{ |\Omega_{\mathrm{CTH}} - \Omega_p(I_p)| \right\}_{p \in P} \right], \tag{4.17}$$

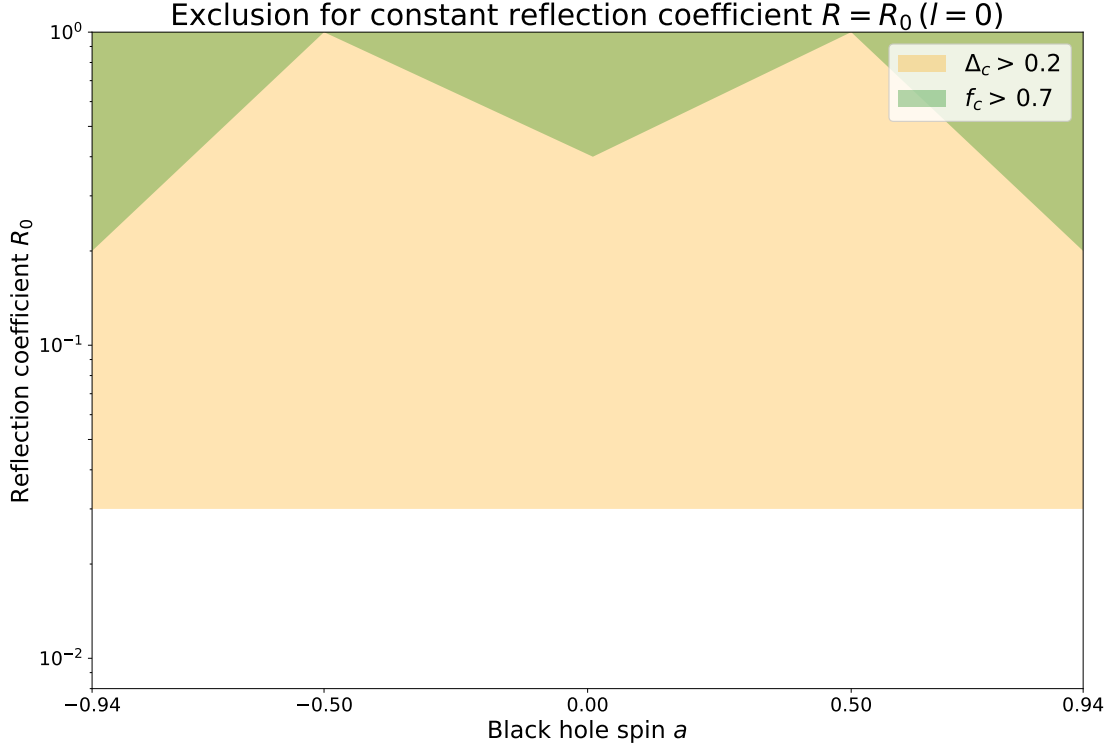where the reflection coefficient R is defined in eq. (2.4).

- We can now add the two images: the unreflected original EHT image and the reflected image created in the previous step. The image outside the photon ring is not modified as those rays do not cross the reflection surface just outside the black hole horizon. The image is normalized such that the sum of all pixel intensities is the same as in the original EHT image. Examples of these images can be found in Figure 2 for $l = 0$ reflection and in Appendix C for higher multipole reflection coefficients $l = 1, 2$.

- Finally we run the image diagnostics discussed in Section 3 to determine if an image is consistent with the diagnostics of the observed image.

Note that this procedure also comes with a few caveats:

- It is a perturbative approach in the sense that we can only constrain small changes to the image as we infer the intensities of the reflected rays from the original (unperturbed) image. However, this assumption is justified as the EHT image is consistent with a Kerr black hole in GR.

- To trace the rays through the accretion disk in the vicinity of the black hole, the medium has to be optically thin. This assumption is consistent with what is known about the accretion disk of M87 [37–39].

- Finally, we assume the reference surface radiates with equal intensity in all directions, *i.e.* the intensity of the reflected rays only depends on where in the reference surface it starts (and not on the direction in which it travels).

A clear advantage of this approach is that we do not have to make additional assumptions or make simulations (on top of those the EHT collaboration is relying on already) about the environment close to the horizon of M87. That being said the ultimate way to constrain these theories of modified gravity described by a reflection coefficient would be to incorporate reflection into general relativistic magnetohydrodynamic (GRMHD) simulations such as those relied upon by the EHT collaboration [41].

5   Results

Figure 3.   Exclusion plot for constant reflection coefficient $l = 0$. The coloured regions are excluded by $\Delta_c > 0.2$ and $f_c > 0.7$, respectively.[9]

We scan over a set of models defined by the possible combinations of the following parameters:

$$l \in [0, 1, 2] \quad \text{and} \quad R_0 \in \left[10^{-3}, 5 \cdot 10^{-3}, 10^{-2}, 3 \cdot 10^{-2}, 7 \cdot 10^{-2}, 10^{-1}, 2 \cdot 10^{-1}, 4 \cdot 10^{-1}, 1\right] . \tag{5.1}$$

One choice of l and R_0 completely defines an effective reflection model according to eq. (2.4). Since the spin of M87 is to date not determined [8] we also scan over a variety of spin parameters, similar to those considered in [41]:

$$a \in [-0.94, -0.5, 0.01, 0.5, 0.94] . \tag{5.2}$$

[8]The exception is that all models with $a = 0$ considered in [41] are excluded by too small jet power.
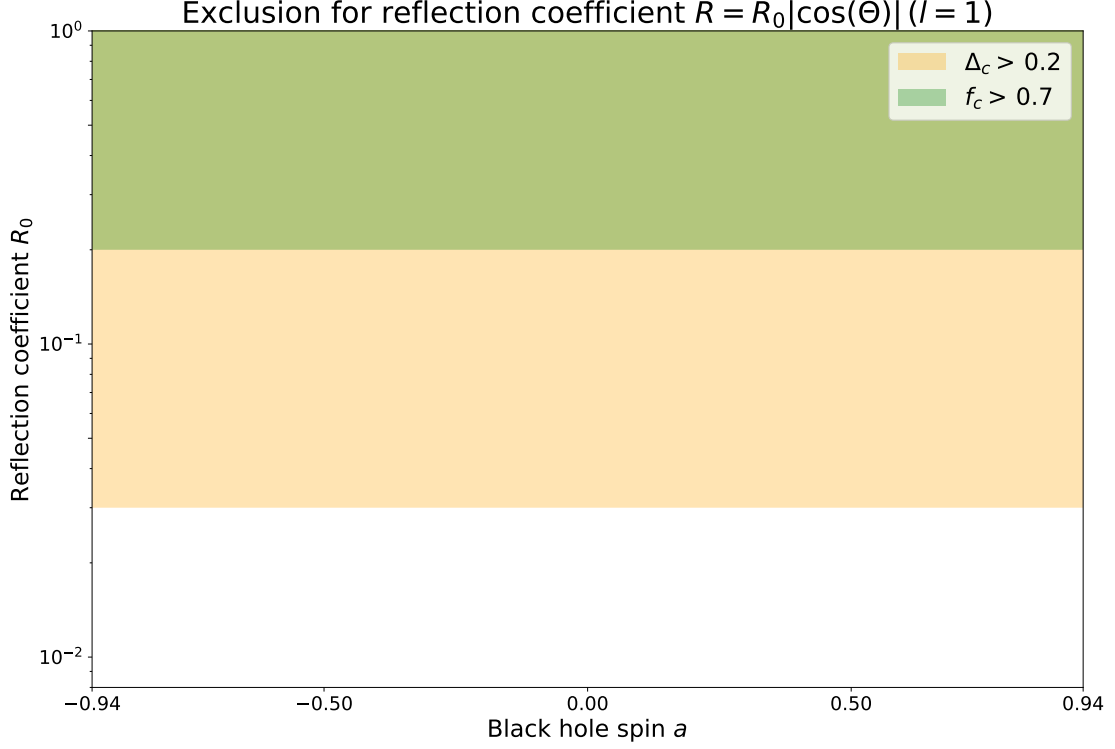
Figure 4. Exclusion plot for the $l = 1$ reflection coefficient. The coloured regions are excluded by $\Delta_c > 0.2$ and $f_c > 0.7$, respectively.[9]

In order to determine if a model is excluded, we evaluate the exclusion criterion formulated at the end of Section 3.1 for every combination of l, R_0 and a, i.e. we test if the image characteristics d, w, η and A are in agreement with the values in the first row of Table 1 and the upper bounds in (3.11) are not violated.

We find that the deviation from circularity Δ_c is the most constraining property, as images with a sizable reflection coefficient tend to increase the brightness towards the center of the image (see e.g. Figure 2) which tends to increase σ_d and hence Δ_c.[9] Another effect of increasing the brightness towards the image center is an increased fractional central brightness f_c. As this quantity is in general less constrained by the EHT observations, the exclusion bounds we find from f_c are about an order of magnitude weaker than those from Δ_c. d, w, η and A are 1-σ compatible with the values in Table 1, except some parameter points at $R_0 = 1$ that are anyways excluded by Δ_c and f_c. Note that the errors we determine for these quantities in Section 3.1 are rather large as we are using a digitized image of the EHT observations. Using the about three times smaller errors on these quantities and stronger bounds on Δ_c and f_c from the original EHT image eq. (3.10), would lead to more competitive bounds than from our digitized image.

Our bounds are summarized in Figures 3 ($l = 0$), 4 ($l = 1$), and 5 ($l = 2$). For the lowest multipole $l = 0$, i.e. angular independent reflection coefficient and the first multipole $l = 1$

---

[9]For the confidence level of exclusions from Δ_c, see Section 7.4 of [42]. [42] does not quote a precise confidence level of their upper bound $\Delta_c < 0.1$ (we use $\Delta_c < 0.2$ as out image quality is worse) but their Figure 18 suggests a confidence level $\gtrsim 90\%$.
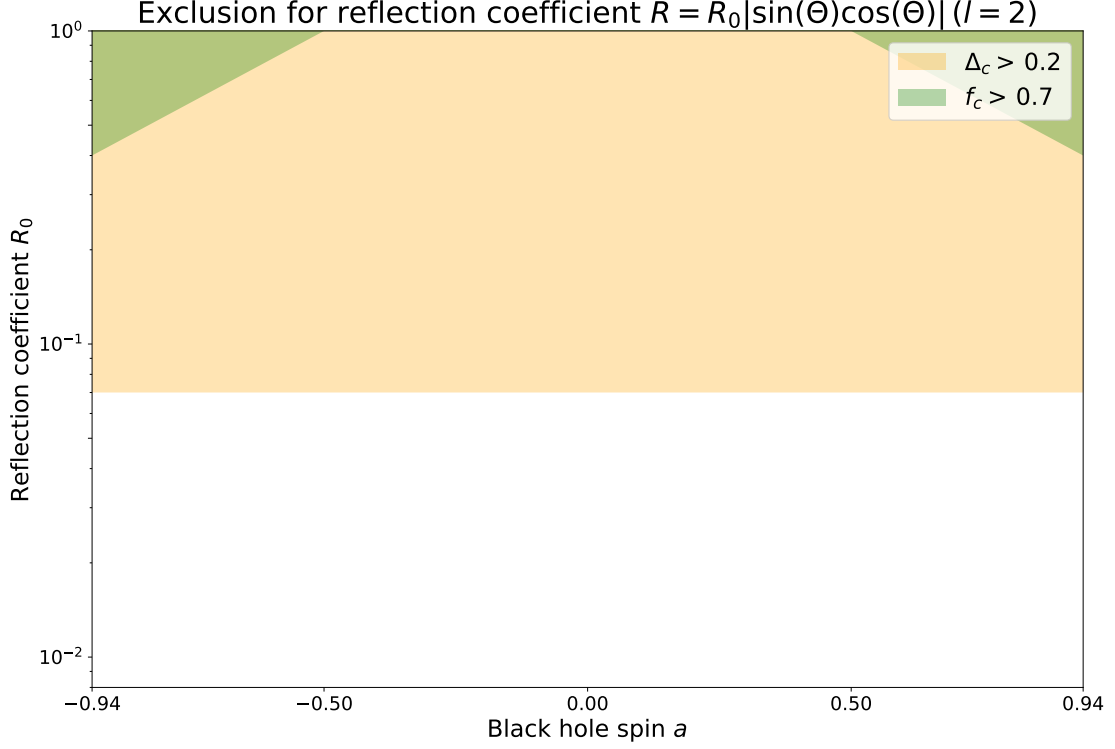
Figure 5. Exclusion plot for the $l = 2$ reflection coefficient. The coloured regions are excluded by $\Delta_c > 0.2$ and $f_c > 0.7$, respectively.[9]

which induces a $\cos\theta$ dependence on the polar angle we find the bound

$$R_0 \lesssim 3 \cdot 10^{-2} \qquad \text{for} \qquad l = 0, 1\,, \tag{5.3}$$

for all possible spins a we consider. For $l = 0$, f_c excludes high spin models $|a| = 0.94$ at $R_0 \lesssim 2 \cdot 10^{-1}$ while for $l = 1$, $R_0 \lesssim 2 \cdot 10^{-1}$ for all values of a.

For $l = 2$, the constraint from Δ_c is weaker at

$$R_0 \lesssim 7 \cdot 10^{-2} \qquad \text{for} \qquad l = 2\,, \tag{5.4}$$

while f_c can only exclude models with $|a| = 0.94$ at $R_0 \lesssim 4 \cdot 10^{-1}$.

Note, that an interesting effect of significant reflection is to focus more brightness inside the photon ring, i.e. into the center of the image. If reflection would be integrated into GRMHD simulations (as opposed to our perturbative treatment) this might lead to smaller ring diameters for a given mass of the supermassive black hole in M87. Since there is already a tension with kinematic measurements preferring a lower mass [46], this could lead to even more stringent constraints on modified gravity theories when these kinematic measurements are taken into account.

6 Conclusions

In this work, we constrained modifications of GR using the EHT observations of M87 [2]. Many interesting extensions of GR [16–18] can be effectively described by a reflection coefficient that specifies how much of an ingoing wave is reflected into an outgoing wave at a surface

close to where the event horizon of a classical black hole would be. The phenomenology of such a reflecting black hole system can be coherently described in a PPEFT framework as discussed in [20].

If the outgoing/reflected wave reaches the detector, in this case the EHT, it modifies the image. We constructed images from modified gravity theories using a ray tracing algorithm and the EHT observations to assign intensities to the reflected rays. While our approach is useful to constrain modified gravity theories that only effect the EHT observations perturbatively - which is justified since the observations are consistent with a Kerr black hole - it would be a natural next step to include reflection into GRMHD simulations that are used by the EHT [41].

We used image diagnostics such as the image rings' deviation from circularity and fractional central brightness to determine if the modified images from reflection are consistent with the EHT observations. We find that we can constrain the RG invariant reflection coefficient R to be less than $1 - 10\%$ with a weak dependence on black hole spin and angular dependence of R. We expect these constrains to improve with the original EHT dataset and GRMHD simulations.

Acknowledgments

A Reflection coefficients and EFT couplings

In this section, we briefly summarize how the RG invariant reflection coefficient R is related to the EFT coupling h_1 of eq. (2.2) and the new-physics RG-invariant length-scale ℓ. For a more in-depth discussion see [20] and [24].

Exotic UV physics that is localized near the horizon affects physics far from the horizon through the changes it makes to the near-horizon boundary condition experienced by any external low-energy 'bulk' fields used to probe the near-horizon regime. In the EFT formalism of refs. [21–23], any particular type of modified near-horizon boundary condition is captured in terms of a surface contribution,

$$S_b = \int_\Sigma \mathrm{d}^3 x \, \mathfrak{L}_b \,, \tag{A.1}$$

to the low-energy effective action whose presence 'tells' the low-energy theory that the boundary condition gets modified.[10] The surface Σ can be anywhere, and the effective couplings in S_b depend on Σ in precisely the way they must to ensure that observables are Σ-independent. For UV physics localized near the would-be horizon the cleanest split between UV and other effects arises if Σ is chosen in the near-horizon regime (but outside the region where the UV

---

[10]This is similar in spirit to the black-hole 'membrane paradigm' [55] used in black-hole astrophysics. One way of thinking about the EFT of [20] is as a theoretical framework in terms of which this paradigm can be derived (including systematic corrections) and adapted to extensions of GR.

physics is important). In this language any freedom of choice in the nature of R appears as the freedom to choose effective couplings within the boundary action S_b.

When the dust settles, physical quantities (like reflection coefficients) are RG invariants in the sense that they depend on effective couplings in a way that is Σ-independent. This implies they can be expressed in terms of RG-invariant characterizations of coupling-constant flow. For the simplest couplings — like that of (2.2) — RG evolution turns out to be labeled by two parameters: an RG-invariant phase, $e^{i\Theta_\star}$, and length scale, ℓ, via a process similar to dimensional transmutation. ℓ and $\Theta_\star$ are RG-invariant parameters that uniquely specify a particular RG flow line, and physically parameterize the relative amplitudes of reflection and absorption as well as any phase shift associated with reflection. Of these only ℓ is needed to determine $|R|$, so this is all that is required in the discussion of the main text.

The physical reflection coefficient R discussed in this work is related to the EFT as

$$R = |R_N R_{\mathrm{UV}}|\,, \tag{A.2}$$

where R_N is present even for a black hole and depends on the grey-body factors and normalization constants of the full mode functions (but not on h_1, ϵ, *etc.*) and is hence not associated with UV physics. The UV physics is all within R_{UV}, which is given by

$$R_{\mathrm{UV}} = \frac{2i\xi + \lambda(\epsilon)}{2i\xi - \lambda(\epsilon)} \left(\frac{\epsilon}{r_s}\right)^{-2i\xi}\,, \tag{A.3}$$

where, for boundary conditions in Kerr spacetime specified at the radial coordinate $r = r_H + \epsilon$ (in Boyer-Lindquist coordinates – see (4.3)),

$$\lambda(\epsilon) = \frac{1}{2\pi\epsilon} \left[\frac{\rho(r,\theta)h_1(\theta,\epsilon)}{\sqrt{\Delta(r)}}\right]_{r=r_H+\epsilon} + 1\,, \tag{A.4}$$

where θ-independent R requires a θ-dependence of $h_1(\theta,\epsilon)$ such that $\lambda(\epsilon)$ is independent of θ. For a Kerr black hole with angular momentum parameter a and mass M, the constant ξ is given as

$$\xi = \left\{\frac{(\omega r_s r_H - am)^2 + is\left[ma\sqrt{r_s^2 - 4a^2} - (r_H^2 - a^2)r_s\omega\right]}{r_s^2 - 4a^2} - \frac{s^2}{4}\right\}^{1/2}\,, \tag{A.5}$$

where $r_s = 2GM$. This depends both on the spin s of the bulk field (where $s = 0$ is a Klein-Gordon scalar, $s = 1$ an electromagnetic field and $s = 2$ a spin-two metric fluctuation), the solutions' energy ω and magnetic quantum number m. In the Schwarzschild ($a \to 0$) limit this becomes

$$\xi \to \omega\, r_s - \frac{is}{2} \qquad \text{(Schwarzschild limit)}\,. \tag{A.6}$$

Since R_{UV} is an RG invariant, the dependence of h_1 on ϵ is precisely such that it cancels the ϵ dependence of the remaining terms on the RHS of eq. (A.3). Alternatively, this can be described by relating R_{UV} to the RG invariant length scale ℓ, defined *e.g.* as the scale for which $\lambda(\epsilon = \ell) = 0$. For the phenomenology explored in this work, only $|R_{\mathrm{UV}}|$ is relevant , and is related to ℓ via

$$|R_{\mathrm{UV}}| = \left(\frac{\ell}{r_s}\right)^{2\,\mathrm{Im}(\xi)}\,. \tag{A.7}$$

B Kerr geodesics

Here, we list the expressions $F^i(x^j, p^j)$ necessary to solve the Kerr geodesic differential equation (4.7):

$$F^r = \frac{1}{2C} \left[\frac{\partial A}{\partial r} \dot{t}^2 - 2\frac{\partial B}{\partial r} \dot{t} p^\phi - \frac{\partial C}{\partial r}(p^r)^2 - 2\frac{\partial C}{\partial \theta} p^\theta p^r + \frac{\partial D}{\partial r}(p^\theta)^2 + \frac{\partial F}{\partial r}(p^\phi)^2 \right],$$

$$F^\theta = \frac{1}{2D} \left[\frac{\partial A}{\partial \theta} \dot{t}^2 - 2\frac{\partial B}{\partial \theta} \dot{t} p^\phi + \frac{\partial C}{\partial \theta}(p^r)^2 - 2\frac{\partial D}{\partial r} p^\theta p^r - \frac{\partial D}{\partial \theta}(p^\theta)^2 + \frac{\partial F}{\partial \theta}(p^\phi)^2 \right], \qquad \text{(B.1)}$$

$$F^\phi = \frac{1}{G} \left[E\frac{\partial H}{\partial r} p^r + E\frac{\partial H}{\partial \theta} p^\theta - \frac{\partial G}{\partial r} p^r p^\phi - \frac{\partial G}{\partial \theta} p^\theta p^\phi \right],$$

where $\dot{t}$ is eliminated via energy conservation

$$\dot{t} = \frac{E}{A} + Hp^\phi, \qquad \text{(B.2)}$$
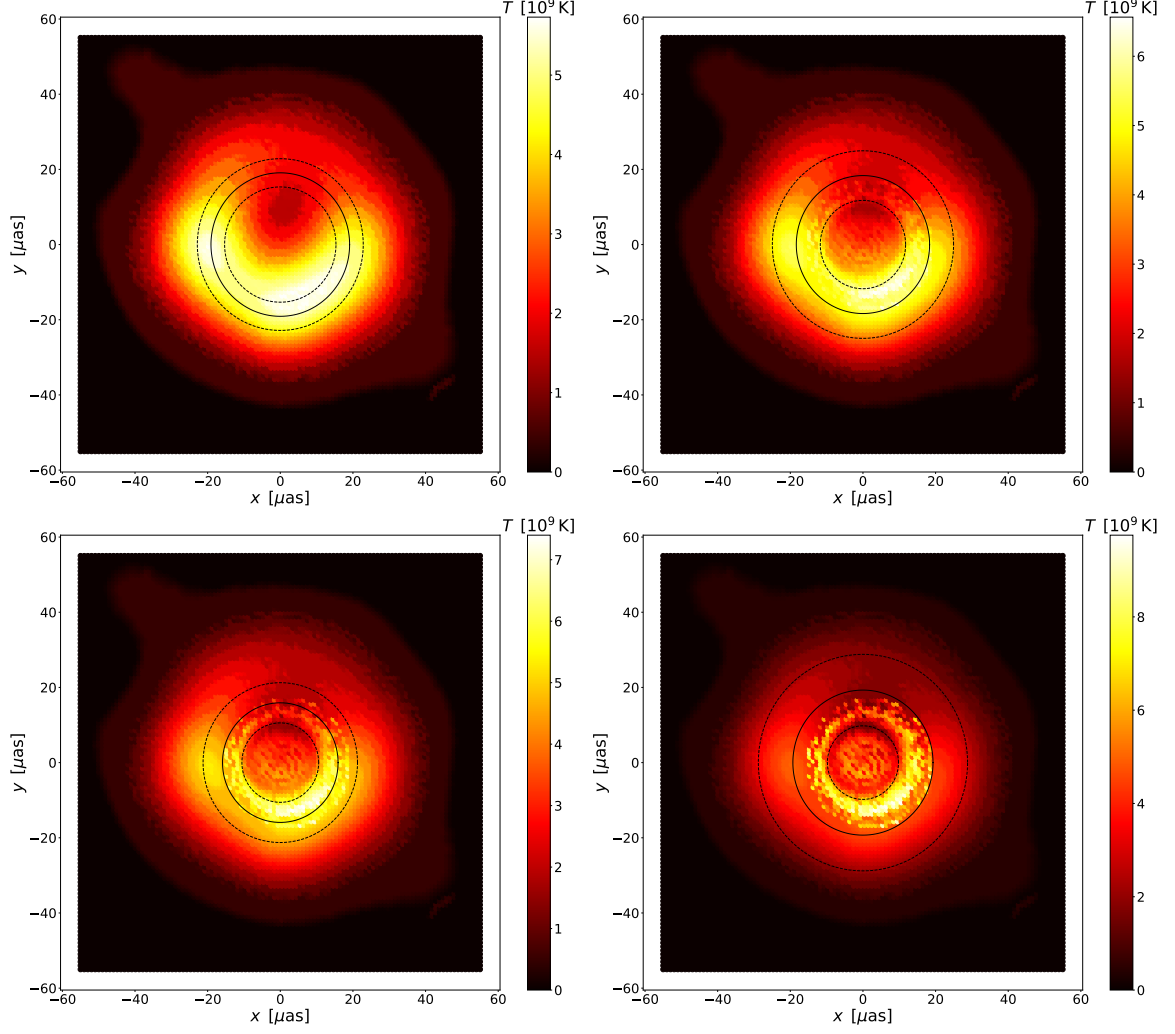
and

$$A = 1 - \frac{r_s r}{\rho^2}, \qquad B = \frac{a\,r\,r_s \sin^2\theta}{\rho^2}, \qquad C = \frac{\rho^2}{\Delta}, \qquad D = \rho^2,$$

$$F = \sin^2\theta \left(r^2 + a^2 + \frac{a\,r\,r_s \sin^2\theta}{\rho^2} \right), \qquad G = F - \frac{B^2}{A} = \sin^2\theta \left(r^2 + a^2 + \frac{a\,r\,r_s \sin^2\theta}{\rho^2 - r\,r_s} \right),$$

$$H = \frac{B}{A} = \frac{a\,r\,r_s \sin^2\theta}{\rho^2 - r\,r_s}.$$

$$\text{(B.3)}$$

C Images for multipoles $l = 1, 2$

Here we show some examples of EHT images for higher multipole reflection coefficients $l = 1$ (Figure 6) and $l = 2$ (Figure 7). The procedure how to create these images is described in Section 4.2.

Figure 6. The added EHT image for $R_0 = 0.01, 0.2, 0.4, 1$ (top left to bottom right) for $l = 1$ and black hole spin $a = 0.94$. The solid black line is half a ring diameter $d/2$ from the image center where d is defined in eq. (3.3). The dashed black lines are at radii $(d - \sigma_d)/2$ and $(d + \sigma_d)/2$, respectively where σ_d is defined in (3.4).
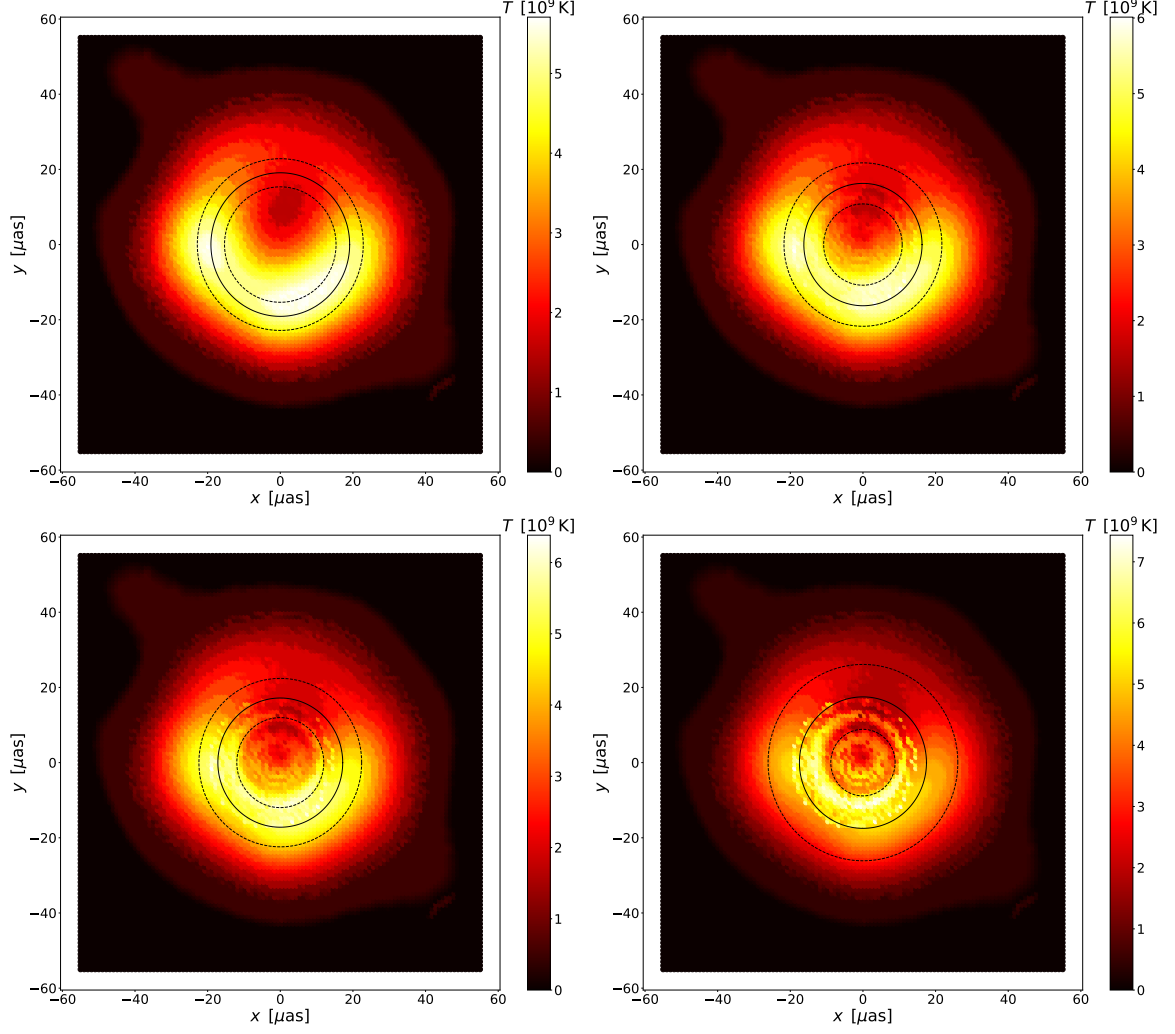
Figure 7. The added EHT image for $R_0 = 0.01, 0.2, 0.4, 1$ (top left to bottom right) for $l = 2$ and black hole spin $a = -0.94$. The solid black line is half a ring diameter $d/2$ from the image center where d is defined in eq. (3.3). The dashed black lines are at radii $(d - \sigma_d)/2$ and $(d + \sigma_d)/2$, respectively where σ_d is defined in (3.4).

References

[1] LIGO Scientific Collaboration and Virgo Collaboration collaboration, *Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs*, *Phys. Rev. X* **9** (2019) 031040.

[2] Event Horizon Telescope collaboration, *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, *Astrophys. J.* **875** (2019) L1 [`1906.11238`].

[3] E. Berti et al., *Testing General Relativity with Present and Future Astrophysical Observations*, *Class. Quant. Grav.* **32** (2015) 243001 [`1501.07274`].

[4] B. S. Sathyaprakash et al., *Extreme Gravity and Fundamental Physics*, `1903.09221`.

[5] C. P. Burgess, *Quantum gravity in everyday life: General relativity as an effective field theory*, *Living Rev. Rel.* **7** (2004) 5 [`gr-qc/0311082`].

[6] J. F. Donoghue, *The effective field theory treatment of quantum gravity*, *AIP Conf. Proc.* **1483** (2012) 73 [`1209.3511`].

[7] W. D. Goldberger, *Les Houches lectures on effective field theories and gravitational radiation*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006*, 2007, `hep-ph/0701129`.

[8] C. P. Burgess, *An Introduction to Effective Field Theory: Thinking Effectively About Hierarchies of Scale*. Cambridge University Press, 2019.

[9] W. D. Goldberger and I. Z. Rothstein, *An Effective field theory of gravity for extended objects*, *Phys. Rev.* **D73** (2006) 104029 [`hep-th/0409156`].

[10] G. Allwright and L. Lehner, *Towards the nonlinear regime in extensions to GR: assessing possible options*, *Class. Quant. Grav.* **36** (2019) 084001 [`1808.07897`].

[11] J. Cayuso, N. Ortiz and L. Lehner, *Fixing extensions to general relativity in the nonlinear regime*, *Phys. Rev.* **D96** (2017) 084043 [`1706.07421`].

[12] M. Okounkova, L. C. Stein, M. A. Scheel and S. A. Teukolsky, *Numerical binary black hole collisions in dynamical Chern-Simons gravity*, *Phys. Rev.* **D100** (2019) 104026 [`1906.08789`].

[13] M. Okounkova, L. C. Stein, J. Moxon, M. A. Scheel and S. A. Teukolsky, *Numerical relativity simulation of GW150914 beyond general relativity*, `1911.02588`.

[14] C. P. Burgess and M. Williams, *Who You Gonna Call? Runaway Ghosts, Higher Derivatives and Time-Dependence in EFTs*, *JHEP* **08** (2014) 074 [`1404.2236`].

[15] A. R. Solomon and M. Trodden, *Higher-derivative operators and effective field theory for general scalar-tensor theories*, *JCAP* **1802** (2018) 031 [`1709.09695`].

[16] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black Holes: Complementarity or Firewalls?*, *JHEP* **02** (2013) 062 [`1207.3123`].

[17] L. Susskind, *The Transfer of Entanglement: The Case for Firewalls*, `1210.2098`.

[18] O. Lunin and S. D. Mathur, *AdS / CFT duality and the black hole information paradox*, *Nucl. Phys.* **B623** (2002) 342 [`hep-th/0109154`].

[19] S. D. Mathur, *The Fuzzball proposal for black holes: An Elementary review*, *Fortsch. Phys.* **53** (2005) 793 [`hep-th/0502050`].

[20] C. P. Burgess, R. Plestid and M. Rummel, *Effective Field Theory of Black Hole Echoes*, *JHEP* **09** (2018) 113 [`1808.00847`].

[21] C. P. Burgess, P. Hayman, M. Williams and L. Zalavari, *Point-Particle Effective Field Theory I: Classical Renormalization and the Inverse-Square Potential*, *JHEP* **04** (2017) 106 [`1612.07313`].

[22] C. P. Burgess, P. Hayman, M. Rummel, M. Williams and L. Zalavari, *Point-Particle Effective Field Theory II: Relativistic Effects and Coulomb/Inverse-Square Competition*, *JHEP* **07** (2017) 072 [1612.07334].

[23] C. P. Burgess, P. Hayman, M. Rummel and L. Zalavari, *Point-Particle Effective Field Theory III: Relativistic Fermions and the Dirac Equation*, *JHEP* **09** (2017) 007 [1706.01063].

[24] R. Plestid, C. P. Burgess and D. H. J. O'Dell, *Fall to the Centre in Atom Traps and Point-Particle EFT for Absorptive Systems*, *JHEP* **08** (2018) 059 [1804.10324].

[25] C. P. Burgess, P. Hayman, M. Rummel and L. Zalavari, *Reduced theoretical error for* $^4He^+$ *spectroscopy*, *Phys. Rev.* **A98** (2018) 052510 [1708.09768].

[26] V. Cardoso, E. Franzin and P. Pani, *Is the gravitational-wave ringdown a probe of the event horizon?*, *Phys. Rev. Lett.* **116** (2016) 171101 [1602.07309].

[27] V. Cardoso, S. Hopper, C. F. B. Macedo, C. Palenzuela and P. Pani, *Gravitational-wave signatures of exotic compact objects and of quantum corrections at the horizon scale*, *Phys. Rev.* **D94** (2016) 084031 [1608.08637].

[28] J. Abedi, H. Dykaar and N. Afshordi, *Echoes from the Abyss: Tentative evidence for Planck-scale structure at black hole horizons*, *Phys. Rev.* **D96** (2017) 082004 [1612.00266].

[29] A. Maselli, S. H. Völkel and K. D. Kokkotas, *Parameter estimation of gravitational wave echoes from exotic compact objects*, *Phys. Rev.* **D96** (2017) 064045 [1708.02217].

[30] G. D'Amico and N. Kaloper, *On Black Hole Echoes*, 1912.05584.

[31] C. Bambi, K. Freese, S. Vagnozzi and L. Visinelli, *Testing the rotational nature of the supermassive object M87* from the circularity and size of its first image*, *Phys. Rev.* **D100** (2019) 044057 [1904.12983].

[32] R. Shaikh and P. S. Joshi, *Can we distinguish black holes from naked singularities by the images of their accretion disks?*, *JCAP* **1910** (2019) 064 [1909.10322].

[33] S. Vagnozzi and L. Visinelli, *Hunting for extra dimensions in the shadow of M87**, *Phys. Rev.* **D100** (2019) 024020 [1905.12421].

[34] I. Banerjee, S. Chakraborty and S. SenGupta, *Silhouette of M87*: A New Window to Peek into the World of Hidden Dimensions*, 1909.09385.

[35] P. V. P. Cunha, C. A. R. Herdeiro and E. Radu, *EHT constraint on the ultralight scalar hair of the M87 supermassive black hole*, 1909.08039.

[36] I. Banerjee, S. Sau and S. SenGupta, *Implications of axionic hair on shadow of M87**, 1911.05385.

[37] R. Narayan and I.-s. Yi, *Advection dominated accretion: A Selfsimilar solution*, *Astrophys. J.* **428** (1994) L13 [astro-ph/9403052].

[38] R. Narayan and I. Yi, *Advection dominated accretion: Underfed black holes and neutron stars*, *Astrophys. J.* **452** (1995) 710 [astro-ph/9411059].

[39] C. S. Reynolds, T. Di Matteo, A. C. Fabian, U. Hwang and C. R. Canizares, *The 'Quiescent' black hole in M87*, *Mon. Not. Roy. Astron. Soc.* **283** (1996) L111 [astro-ph/9610097].

[40] Event Horizon Telescope collaboration, *First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole*, *Astrophys. J.* **875** (2019) L4 [1906.11241].

[41] Event Horizon Telescope collaboration, *First M87 Event Horizon Telescope Results. V. Physical Origin of the Asymmetric Ring*, *Astrophys. J.* **875** (2019) L5 [1906.11242].

[42] Event Horizon Telescope collaboration, *First M87 Event Horizon Telescope Results. VI. The Shadow and Mass of the Central Black Hole*, *Astrophys. J.* **875** (2019) L6 [1906.11243].

[43] O. Lunin and S. D. Mathur, *Statistical interpretation of Bekenstein entropy for systems with a stretched horizon*, *Phys. Rev. Lett.* **88** (2002) 211303 [`hep-th/0202072`].

[44] V. Cardoso and P. Pani, *Testing the nature of dark compact objects: a status report*, *Living Rev. Rel.* **22** (2019) 4 [`1904.05363`].

[45] M. Cantiello, J. P. Blakeslee, L. Ferrarese, P. Côté, J. C. Roediger, G. Raimondo et al., *The Next Generation Virgo Cluster Survey (NGVS). XVIII. Measurement and Calibration of Surface Brightness Fluctuation Distances for Bright Galaxies in Virgo (and Beyond)*, *ApJ* **856** (2018) 126 [`1802.05526`].

[46] J. L. Walsh, A. J. Barth, L. C. Ho and M. Sarzi, *The M87 Black Hole Mass from Gas-dynamical Models of Space Telescope Imaging Spectrograph Observations*, *Astrophys. J.* **770** (2013) 86 [`1304.7273`].

[47] L. E. Kidder, M. A. Scheel, S. A. Teukolsky, E. D. Carlson and G. B. Cook, *Black hole evolution by spectral methods*, *Phys. Rev.* **D62** (2000) 084032 [`gr-qc/0005056`].

[48] F. H. Vincent, T. Paumard, E. Gourgoulhon and G. Perrin, *GYOTO: a new general relativistic ray-tracing code*, *Class. Quant. Grav.* **28** (2011) 225011 [`1109.4769`].

[49] T. Müller, *GeoViS—Relativistic ray tracing in four-dimensional spacetimes.*, *Computer Physics Communications* (2014) 2301.

[50] A. Bohn, W. Throwe, F. Hébert, K. Henriksson, D. Bunandar, M. A. Scheel et al., *What does a binary black hole merger look like?*, *Class. Quant. Grav.* **32** (2015) 065002 [`1410.7775`].

[51] A. Riazuelo, *Seeing relativity-I: Ray tracing in a Schwarzschild metric to explore the maximal analytic extension of the metric and making a proper rendering of the stars*, *Int. J. Mod. Phys.* **D28** (2018) 1950042 [`1511.06025`].

[52] P. V. P. Cunha and C. A. R. Herdeiro, *Shadows and strong gravitational lensing: a brief review*, *Gen. Rel. Grav.* **50** (2018) 42 [`1801.00860`].

[53] S. Chandrasekhar, *The mathematical theory of black holes*, in *Oxford, UK: Clarendon (1992) 646 p., Oxford, UK: Clarendon (1985) 646 P.*, 1985.

[54] R. Craig Walker, P. E. Hardee, F. B. Davies, C. Ly and W. Junor, *The Structure and Dynamics of the Subparsec Jet in M87 Based on 50 VLBA Observations over 17 Years at 43 GHz*, *Astrophys. J.* **855** (2018) 128 [`1802.06166`].

[55] K. S. Thorne, R. H. Price and D. A. Macdonald, eds., *Black Holes: The Membrane Paradigm.* 1986.