Imaging black holes: past, present and future

Heino Falcke$^{1,2}$
$^1$ Department of Astrophysics/IMAPP, Radboud University, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
$^2$ ASTRON, Dwingeloo, The Netherlands
E-mail: h.falcke@astro.ru.nl

Abstract. This paper briefly reviews past, current, and future efforts to image black holes. Black holes seem like mystical objects, but they are an integral part of current astrophysics and are at the center of attempts to unify quantum physics and general relativity. Yet, nobody has ever seen a black hole. What do they look like? Initially, this question seemed more of an academic nature. However, this has changed over the past two decades. Observations and theoretical considerations suggest that the supermassive black hole, Sgr A*, in the center of our Milky Way is surrounded by a compact, foggy emission region radiating at and above 230 GHz. It has been predicted that the event horizon of Sgr A* should cast its shadow onto that emission region, which could be detectable with a global VLBI array of radio telescopes. In contrast to earlier pictures of black holes, that dark feature is not supposed to be due to a hole in the accretion flow, but would represent a true negative image of the event horizon. Currently, the global Event Horizon Telescope consortium is attempting to make such an image. In the future those images could be improved by adding more telescopes to the array, in particular at high sites in Africa. Ultimately, a space array at THz frequencies, the Event Horizon Imager, could produce much more detailed images of black holes. In combination with numerical simulations and precise measurements of the orbits of stars – ideally also of pulsars – these images will allow us to study black holes with unprecedented precision.

1. Introduction

When the name black hole was coined in the sixties – allegedly a term picked up by science journalist Ann Ewing and later popularized by John A. Wheeler$^1$ – it sparked the imagination: not a dull mathematical, abstract object in a theory of gravity but a somewhat mythical, perhaps scary object with an intriguing visual appearance — something seemingly invisible that nonetheless exists. Ever since, there have been many depictions of black holes in the scientific literature, in the arts, and in the public media — perhaps even more so in the latter.

Ironically, the reason why black holes became fashionable was not because they were black, but because they were bright. This was related to the discovery of quasars more than 50 years ago [see, e.g. 57]. Quasars are extremely luminous and simple arguments could show that the deep potential well of black holes could in principle produce that energy.

A falling test mass $m$ at radius $R$ from a black hole mass $M_*$ woul have a Newtonian potential of $E = GM_*/R$, where $G$ is the gravitational constant. For a constant mass inflow rate $\dot{m}$ the available energy per time $\dot{E}$ can be converted into a luminosity $L$ with an assumed efficiency $\eta$ of order 10% to yield $L \sim \eta \dot{E} = \eta GM_*/R \dot{m} = \frac{1}{2} \eta \left( \frac{R}{R_*} \right) \dot{m} c^2$, where $c$ is the speed of light and $R$ is

$^1$ http://www.worldwidewords.org/topicalwords/tw-bla1.htm
assumed to be of order the Schwarzschild radius, $R_S = \frac{2GM}{c^2}$. For $\dot{m} \sim 1M_\odot$/yr the black hole luminosity $L$ can be of order $10^{46}$ erg/sec for black holes some hundred million solar masses – comparable to inferred quasar luminosities.

Not surprisingly, the topic of early research on black holes and active galactic nuclei (AGN) then mainly focused on the overall spectral energy distribution (SED) of luminous AGN. Because of their large distances, imaging a black hole on scales of the event horizon seemed out of reach and more like a purely academic exercise. The only resolvable structure of quasars was seen in their radio emission that later was identified as coming from relativistic plasma jets emerging from near the putative black holes [e.g., 7, 22].

2. Black hole images: early predictions

So, what does a black hole really look like? The first attempt I am aware of dates back to [11], where the authors calculate the appearance of a star in a tight orbit (3 and 20 Schwarzschild radii) around a Kerr black hole (Fig. 1a). The apparent orbit already traces out the basic structures later also seen by more complete accretion disk simulations. However, what we see in these images is not an image of the event horizon at all, it is actually the lensed image of the stellar orbit and true for any gravitating object with or without an event horizon.

The paper that actually caught my attention first on this topic was by [4], published in

Cunningham & Bardeen also refer to Campbell & Matzner [9], doing a similar calculation for a Schwarzschild black hole, but neither motivation nor results seem particularly relevant here.
a conference proceedings, where it seemed hardly noticed and cited until then\(^3\). In this paper Bardeen calculates the appearance of an extended stellar disk behind a Kerr black hole (Fig. 1c). A similar calculation was later presented by \([36]\) for an extended plane wave source at a large distance, e.g., a star in a binary system or simply a large flash light shining at the black hole from behind.

Here an important effect is noticeable, which is related to the photon orbit. The photon orbit is an orbit where light can go around black holes on a closed loop. In practice this means that light entering at or around the photon orbit, will experience a sharp divide: if the light ray is a little bit too close it will disappear behind the event horizon, if it is a little bit outside, it will be able to escape to infinity. The consequence is indeed a sharp image of the event horizon.

In the same paper Luminet also presents his famous image (Fig. 1b). The figure is hailed as the first ever computer-generated image of a black hole. However, as in the case of \([11]\), the dark structure traced out in the center is not actually an image of the black hole itself, but is a lensed image of the inner edge of the accretion disk. Consequently, the image depends crucially on the assumption of where the disk ends. The standard view at the time \([49, 50]\) was that the disk terminated and stopped radiating at the innermost stable circular radius (ISCO, then often called marginally stable radius \(r_{\text{ms}}\)). The same is true for later efforts \([59, \text{Fig. 1d}]\). Strictly speaking we see therefore in these images mainly a hole in an accretion disk rather than the black hole itself.

### 3. The Galactic Center: where dreams come true

A region of high hopes for black hole research had been the center of our own Galaxy \([37]\). The compact radio source, Sgr A*, found by Balick & Brown \([3]\), was often compared to radio cores in AGN, and hence considered to mark the supermassive black hole in our Milky Way. Near-infrared measurements of stellar orbits \([14, 26]\), revealing stars moving with up to 10,000 km/s around this radio source, and very-long baseline interferometry (VLBI) observations \([54]\), have confirmed that indeed there is a dark mass of about 4.3 million solar masses \([27]\) within a few Schwarzschild radii of Sgr A* \([\text{see 25, 20, for reviews}]^4\).

The nature of this radio emission was long under debate \([41]\). In the early nineties, we developed a model to describe Sgr A* as a starving black hole with a low-power radio jet, fed by a highly sub-Eddington accretion flow \([18, 19]\). A competing model was to explain all the radio emission from a radiatively inefficient accretion flow (RIAF) \([48]\) at much higher accretion rates. Today we often combine those elements \([60, 45]\) and keep the low accretion rate now demanded by polarization observations of Sgr A* \([6, 38]\).

A prediction of the jet model was that the radio emission should come from closer to the black hole the higher the frequency of the emission was, meaning that the submm-wave bump \([42]\) in the spectrum should come from a compact emission region near the event horizon. One of the first multi-telescope and multi-wavelength campaigns for Sgr A*, involving four telescopes, confirmed that picture \([16]\). Simple theory also predicted that the emission should become optically thin at and above 230 GHz. Hence, we proposed in the same paper that one should be able to image the event horizon against this background with submm-wave VLBI observations.

Since the original Bardeen picture only considered emission from far behind the black hole, we then calculated the appearance of a black hole surrounded by an optically thin emission region \([21, \text{Fig. 1e}]\), which was more applicable to Sgr A*. Indeed, we found a dark region of about some 5 Schwarzschild radii in diameter, surrounded by a bright ring of emission. That dark region is not just a sharp silhouette as sometimes implied. Interstellar scattering of radio waves and a blend of emission in front of and behind the black hole will give the feature a more

---

\(^3\) Intriguingly, this is the book, where the famous paper by Novikov & Thorne \([49]\) was published and which has been cited more than 500 times by now.

\(^4\) For parameters of Sgr A* used here, see the overview in Table 1 in Falcke & Markoff (2013).
diffuse appearance. Hence, we called it the “shadow” of the event horizon. Many papers now exist that calculate black holes shadows, also in non-standard theories of gravity, suggesting its usefulness for testing GR [see 20, 30, for reviews].

Simulations are now becoming much more sophisticated and are done using numerical general-relativistic magnetohydrodynamic simulations (GRMHD) with ray tracing to take the astrophysical model into account. This was pioneered by [47], initially for a pure disk model, but now also including jet emission [see 44, for a recent review]. The new BHAC code [51] now also allows one to perform these simulations in arbitrary parametrized space times [32]. In Moscibrodzka & Falcke [45] and Moscibrodzka et al. [46] we showed that the original simple analytical jet model does in fact hold up in GRMHD simulations, reproduces the observational characteristics, and produces radio mm-wave emission close to the event horizon — just as needed for shadow imaging.

The diameter of the shadow would be of order $10 GM_c c^{-2}$, i.e. 63 million km for $M_c = 4.3 \times 10^6 M_\odot$, corresponding to 51 microarcseconds ($\mu$as) for a distance of 8.3 kpc to the Galactic Center. The resolution of a radio interferometer is given by the ratio of observing wavelength and telescope separation $\lambda/D$. For global interferometers $D$ reaches typical values of 15,000 to 6,000 km, giving one a resolution between 18 to 50 $\mu$as at 230 GHz (i.e., $\lambda = 1.3\text{mm}$) — just large enough to start resolving the shadow. Moreover, the scattering effect decreases with $\lambda^2$, also approaching 22 $\mu$as at $\lambda = 1.3\text{mm}$. Hence, 230 GHz becomes a ‘magical’ observing frequency, where interstellar scattering, optical depth in the source, global VLBI resolution, and atmospheric transmission make this experiment just possible. At higher frequencies the earth’s atmosphere becomes too hostile and at lower frequencies the source is scatter broadened and transparent. Hence, we can thank God, that the earth is just large enough and at the right location in our Galaxy, to let us to not only see the light, but also the shadow of a black hole.

4. The experimental development: from the past to the present

The first successful 1.3 mm VLBI experiment measuring Sgr A* was a single baseline observation between the IRAM telescopes on Pico Veleta (Spain) and Plateau de Bure (France) led by the MPIfR in Bonn [33]. The idea of imaging the event horizon in Sgr A* was first more widely discussed – though somewhat skeptically – at the Galactic Center conference in Tucson 1998 [e.g., see transcript of VLBI discussions in 61], followed by more discussions at other meetings. In 2004, we organized an informal evening session at the 30th anniversary meeting on the occasion of the discovery of Sgr A* [6], where the community was finally ready to embrace and support such an experiment. The three speakers at the meeting, G. Bower, S. Doeleman and myself, started a series of telecons, further developing the idea of a global physics-like collaboration for black hole imaging. MIT Haystack observatory kept improving digital VLBI hardware for 3 and 1mm observing and proceeded further in a small ad hoc collaboration [12]. A little earlier, regular 3mm observing had already been made possible via a joint memorandum of understanding between multiple observatories, forming the Global mm-VLBI Array (GMVA, 2003), coordinated by the MPIfR Bonn.

Significant progress was made at lower frequencies. Bower et al. [5] were able to measure the intrinsic size at 7mm and 13 mm, showing for the first time that indeed Sgr A* decreases in size towards higher frequencies as predicted and that it would approach event horizon scales if extrapolated. This trend was soon confirmed by better VLBI measurements at 3mm [56].

5 The observation was at low signal-to-noise and the size seemed much larger than the event horizon, hence, did not get too much attention. However, at the time the black hole mass was thought to be much smaller than it is now and the size expressed in Schwarzschild radii seemed larger than the shadow size. In fact, the measured size in $\mu$as at 230 GHz is consistent with current measurements within the error bars.

6 http://www.aoc.nrao.edu/~gcnews/GCconfs/SgrAstar30 — see small print after “Poster Contributions”.

7 In fact, expressed in the form of a vote and a signed declaration.
Doeleman et al. [12] published the first three-baseline experiment, demonstrating a source size of $37^{+15}_{-10}\mu\text{as}$, i.e. smaller than the shadow size. Within the errors the results were consistent with the Krichbaum et al. [33] results, but more robust and with significantly smaller error bars. The triangle SMTO (Arizona) – CARMA (California) – JCMT/SMA (Hawaii) had a number of productive observing runs, showing asymmetric structure and evidence for homogenous magnetic fields [24, 31]. Also, a four-baseline experiment, involving the APEX telescope, succeeded, revealing evidence for more event-horizon scale structure [34, Lu et al. 2018, in prep.].

Community white papers to support mm-VLBI efforts were published in the US, Europe, and Asia [17, 23, 58, 1] and mm-VLBI was included in the European Science Vision for astronomy in 2007 and in the US Astro2010 decadal review. In the summer of 2017 a formal memorandum of understanding was signed by 13 stakeholder institutions with about 150 individual scientists to form the Event Horizon Telescope Consortium (EHTC), which currently is conducting observations with 8 telescopes (IRAM 30m, JCMT, SMA, SMTO, SPT, LMT, ALMA, APEX) on six mountains\(^9\). Besides outfitting telescopes with proper VLBI-equipment, getting ALMA VLBI-ready had been another important step [40]. The EHTC experiments are ongoing and one has to await further data analysis, before any results can be discussed.

Despite the high bandwidth and some large telescopes, the limited number of simultaneous baselines will make imaging still difficult. Simulations by Lu et al. [35] suggest that images made during a single day suffer significant uncertainties from source variability. Substructures and variations in the scattering screen will amplify this uncertainty further. Hence, multiple observing epochs may need to be added for a good image.

In the end, one will have to fit astrophysical shadow images to the experimental data [e.g., 8] to derive black hole parameters and to verify GR. However, the variations in those models lead to an unknown systematic error that still needs to be quantified better. Assuming GR is correct, shadow imaging could lead to an accuracy of up to 10 % [53] in confirming the null-hypothesis that a GR event horizon exists. If, however, GR itself is to be tested against alternative models, it may not be that easy to discriminate between GR shadow images and non-GR images [43].

Hence, like in cosmology, where multiple experiments are combined to provide, e.g., constraints on the nature of dark energy, it may also become important to add complementary information to the VLBI data. Psaltis, Kramer, & Wex (2016) show that measuring a pulsar in the Galactic center would be ideal for this. Indeed, after decades of searches, a first pulsar has been found [13] but closer and older ones are needed to measure the spacetime around Sgr A* accurately. The SKA, but also ALMA, could detect more pulsars in the future. Moreover, more and better stellar orbits will also constrain spacetime in a significant way. Here ESO’s new near-infrared interferometer, GRAVITY, will make a big impact [29].

5. The future: Africa mm-wave telescope and space VLBI
The global millimeter VLBI array will further expand. ASIAA plans to add the Greenland telescope (GLT) for observations of M87 [39] and the BlackHoleCam team [28] together with IRAM, is preparing to phase up\(^{10}\) the NOEMA array (Plateau de Bure, France) for VLBI. Also, the University of Arizona Kitt Peak telescope and the LLAMA dish (Brazil, Argentina) may be added soon.

Moreover, we are currently investigating possibilities to add an African mm-wave telescope (AMT) to the array. The Gamsberg mountain in Namibia is a potential site [2]. First simulations [55, see also Fig. 2] indicate that an Africa telescope would add more robustness and quality to

---

8 http://www.eso.org/public/news/eso0744/

9 The term “Event Horizon Telescope” came out of a coffee table discussion at the AAS meeting in Long Beach in 2009.

10 Phasing-up an array means to add the signals from individual antennas in an array in phase, such that the telescopes act collectively as a large single dish.
the imaging. If multiple epochs are added, the AMT – despite being smaller – would become as important as ALMA for the image quality. In the case of long observations and high signal-to-noise ratio, the distribution of the telescopes is more important than the sensitivity of individual telescopes. Of course, more telescopes in Africa, e.g., near Kilimanjaro in Tanzania, or on the Drakensberg in Lesotho, would improve the quality even further.

Figure 2 Left: Average image of the time variable black hole “movie” used in Lu et al. (2016), b) same image convolved with scattering kernel to account for interstellar scattering, c) synthetic image recovered with an array including SMA, SMT, LMT, ALMA, IRAM 30m, NOEMA, SPT, and an Africa telescope (AMT) using the procedure described in Lu et al. and averaging 8 epochs.

Of course, as always, space is the final frontier. Roelofs et al. [55] performed simulations using a two-element space-to-space interferometer, the Event Horizon Imager (EHI), consisting of two spacecraft at about 15,000 km. The telescopes are separated by some 50 km in altitude (Fig. 3) and observe at 690 GHz. At these frequencies interstellar scattering is almost negligible (∼ 2μas), the source will be very optically thin, and a ground-based VLBI array is unfeasible due to strong and variable tropospheric absorption. Having only a single space-to-space baseline, rather than a fleet of spacecraft, greatly reduces the complexity.

While orbiting the earth, the spacecraft slowly drift apart due to different orbital speeds. Eventually, the connecting baseline will have assumed all orientations and separations, giving one a spiral pattern in the uv-plane — the Fourier plane of the image. A filled uv-plane greatly improves image reconstruction in interferometric measurements. For realistic receiver noise and source fluxes, two 6 m dishes could obtain a very detailed picture of the black hole shadow (Fig. 3) with 4 μas resolution, when integrating over several months or even years — thereby also integrating over source variability. The constant features are those produced by GR.

6. Conclusion
All theoretical and experimental data support the idea that imaging the shadow of the black hole in the Galactic Center (and probably also in M87) should be possible with global radio interferometers at 230 GHz and higher. For testing general relativity, it is crucial that the emission region indeed extends well inside the ISCO, otherwise we would mainly see a hole in the accretion flow, rather than a black hole shadow. Hence, understanding the astrophysical model remains important. Current GRMHD models with a radiatively inefficient accretion flow and a jet outflow already explain the data for Sgr A* and M87* well. The newly formed Event Horizon Telescope Consortium is now attempting to make the first such images. Simulations suggest that multiple epochs may be necessary to obtain a robust image. If successful they could verify the prediction of an event horizon to within 10% accuracy, assuming GR is correct.
Testing alternative theories of gravity would benefit from additional inputs, such as better stellar orbits and ideally, a pulsar around the black hole. Images can be improved by adding additional telescopes to the global array. An African mm-wave telescope (AMT) would add crucial new baselines, producing better and more robust images. Ultimately, a space interferometer, such as the Event Horizon Imager operating at THz frequencies, could produce a quantum leap in black hole imaging. This may be possible even with just two spacecraft. Hence, it is probably more a question of when, rather than whether, the first good images of black holes will arrive. Since most supermassive black holes stay around for a long time, those images can and will become better with time. This provides us with the golden opportunity to conduct more and more detailed tests of gravity in its most extreme limit.

Acknowledgments
I am grateful for comments from F. Roelofs, M. Moscibrodzka, A. Roy, N. Nagar, G. Crew, and other colleagues. The author is supported by the European Research Council (ERC) Synergy Grant BlackHoleCam (Grant 610058), which is jointly led by the author together with L. Rezzolla (Univ. Frankfurt), and M. Kramer (MPIfR Bonn).

References


[51] Porth, O., Olives, H., Mizuno, Y., et al. 2017, Computational Astrophysics and Cosmology, 4, 1