

A Survey on Black Hole Image Reconstruction

Tianao Li¹, Yuxuan Li², Zhiyao Li²

¹Department of Computer Science, Northwestern University

²Department of Electrical and Computer Engineering, Northwestern University

June 6, 2024

Abstract

The Event Horizon Telescope (EHT) collaboration's groundbreaking image of the M87 black hole ignited a new era in black hole imaging. This endeavor requires sophisticated image reconstruction algorithms that combine data from global telescopes. However, reconstructing black holes presents a unique challenge: it's an ill-posed inverse problem. Unlike conventional imaging, there are infinitely many possible solutions that fit the collected data. Traditionally, researchers address this by incorporating prior knowledge, which can be unreliable for groundbreaking discoveries like black holes. Additionally, quantifying the uncertainty in the reconstructed image is crucial for astronomers to assess its reliability. These two aspects, uncertainty quantification and dependence on explicit priors, represent the major hurdles in black hole imaging.

This survey explores recent advancements that address these challenges. We focus on methods that reconstruct black hole images with uncertainty without relying on explicit priors. We'll explore how these methods leverage generative models within a variational Bayesian framework to overcome the limitations of traditional approaches. We'll also delve into the limitations of these recent methods and propose promising avenues for future research.

1 Background

In this section, we will introduce the scientific problem of imaging the black hole, why it is important in astrophysics, and why it is a difficult problem. We will formalize the problem mathematically and point out the major challenges.

1.1 Imaging the Black Hole

Black holes are fundamental to understanding the dynamics and evolution of galaxies. By studying the region surrounding a black hole, particularly the event horizon, scientists can gain critical insights into the processes of accretion, where matter spirals into the black hole, and jet formation, where particles are ejected at relativistic speeds. These phenomena play essential roles in regulating star formation and distributing matter across the universe, influencing the overall structure and evolution of galaxies.

The importance of imaging black holes extends far beyond the realms of theoretical physics and astronomy. High-resolution images of black holes can uncover the underlying mechanisms behind high-energy astrophysical processes, such as the generation of relativistic jets, which can have profound impacts on entire galaxies. By providing direct visual evidence of these processes, such images help to refine our theoretical models of black hole growth and evolution, leading to a more precise and comprehensive understanding of the history and structure of the universe.

Moreover, black hole imaging can validate and expand upon current physical theories. For example, the event horizon of a black hole provides a distinct setting to examine the principles of general relativity in extreme conditions, potentially revealing new physics beyond our current

understanding. The detailed study of black holes also contributes to broader fields of research, such as quantum mechanics and cosmology, by challenging and enhancing our knowledge of the fundamental laws that govern the universe.

In 2019, the Event Horizon Telescope (EHT) reached a historic milestone. They captured an image of the supermassive black hole for the very first time in human history [1]. This black hole is located at the center of the galaxy M87, known as M87*. This groundbreaking achievement highlighted the transformative potential of high-resolution black hole imaging to test the predictions of general relativity and explore the behavior of matter under the influence of extreme gravitational forces [2].

Imaging black holes is one of the most formidable challenges in modern astrophysics. The boundary of a black hole, known as the event horizon, is the point beyond which nothing, not even light, can escape from the black hole's gravity. When observed from Earth, this boundary appears extraordinarily small. Achieving the necessary angular resolution to image this region is akin to being able to see a grain of sand on the moon from Earth—a task of immense precision and difficulty. The angular resolution of a telescope system is determined by diffraction $\theta \sim \frac{\lambda}{D}$ [3], where λ represents the wavelength of the observed signal, and D represents the diameter of telescope's aperture. The angular resolution required for imaging a black hole is at least an order of magnitude finer than that used for previous astronomical observations. For example, if we consider the black hole in our Milky Way, Sag A*, imaging the plasma encircling that black hole necessitates resolving an angle as small as $\sim 3 \times 10^{-5}$ arcsec at a wavelength of 1.3 mm [4]. Achieving such fine resolution demands a telescope with a diameter of $\sim 1 \times 10^4$ km, which is roughly the size of the Earth. However, it is unrealistic to build an Earth-sized telescope.

1.2 Very Long Baseline Interferometry

Although building a single telescope with such an immense diameter is unfeasible, we can emulate the effect by gathering data concurrently from multiple telescopes positioned across the Earth. This technique is **very long baseline interferometry (VLBI)** [3]. VLBI is an advanced astronomical method that allows for the imaging of celestial objects with exceptionally high angular resolving power through utilizing a network of radio receiver dispersed over extensive geographical areas. This technique effectively creates a large imaginary telescope whose aperture is matching the maximum distance between the telescopes in the network, achieving resolutions far beyond what a single telescope could provide.

VLBI works by simultaneously observing a celestial object with multiple radio telescopes positioned in various locations around the world. Each telescope in the array collects radio signals emitted by the object, and these signals are recorded along with highly accurate time stamps provided by atomic clocks. The time-stamped data from all the telescopes are then brought together for processing and analysis. Each pair of telescopes in the array forms a baseline, which is the distance between them. These baselines act as the arms of an interferometer. The primary data product of a VLBI observation is called **visibility**, which measures the correlation between the signals received by two different telescopes. These visibilities are samples of the 2D Fourier space, termed UV plane, of the source's brightness pattern, providing data on the frequency components of the observed object.

The collection of visibilities from different baselines form samples from the UV plane. However, due to the restricted number of telescopes on Earth, as well as their fixed positions, the UV plane is usually sparsely sampled [3]. This sparse coverage results in incomplete information, which poses significant challenges for image reconstruction. As the Earth rotates, the relative positions of the telescopes change, allowing different parts of the UV plane to be sampled over time. This rotation helps to improve the coverage but does not completely fill the UV plane.

After the observations, the data from each telescope are sent to a central processing facility where their correlation are calculated. The correlation process involves cross-matching the signals from each pair of telescopes, taking into account the time delays introduced by the different paths

the signals travel. This step requires precise synchronization, often achieved using highly accurate atomic clocks like hydrogen masers. The correlated data are subsequently calibrated to account for various systematic effects, including atmospheric distortions, instrumental errors, and geometric delays. Atmospheric distortions are particularly challenging at higher frequencies and shorter wavelengths, as variations in the atmosphere's moisture content and turbulence can significantly affect the radio signals.

The final step in VLBI is image reconstruction from calibrated visibilities. This process involves solving an inverse problem, aiming to identify the most probable image that matches the observed visibilities. Due to the sparse sampling of the UV plane, this process is highly ill-posed, which means there are numerous potential images that could fit the data.

The VLBI forward model is shown here. This forward model is used to describe how the signals collected by pairs of telescopes (referred to as baselines) are related to the actual image of the observed astronomical object [5]:

$$\Gamma_{a,b}^t = g_a^t g_b^{t*} e^{i(\phi_a^t - \phi_b^t)} F_{a,b}^t(\mathbf{x}) + \eta_{a,b}^t. \quad (1)$$

Here, a and b denote a pair of telescopes. $\Gamma_{a,b}^t$ represents the observed visibility data for the baseline. $F_{a,b}^t(\mathbf{x})$ represents the Fourier transform of the actual sky brightness distribution \mathbf{x} , sampled at the spatial frequency corresponding to the baseline. This Fourier transform connects the image space to the visibility space. g_a^t and g_b^{t*} are time-dependent telescope-based gain error. They represent the errors in instrument calibrations. $e^{i(\phi_a^t - \phi_b^t)}$ are time-dependent telescope-based phase errors, usually caused by atmospheric turbulence. $\eta_{a,b}^t$ represents thermal noise (usually a Gaussian) in the observed visibility data for the baseline. These noises make the VLBI forward model non-convex, which requires constraints in inverse problem-solving.

In short, VLBI image reconstruction faces challenges due to its non-convex forward model, sparse UV space coverage, and the presence of diverse noise sources, hindering the recovery of the true object's image.

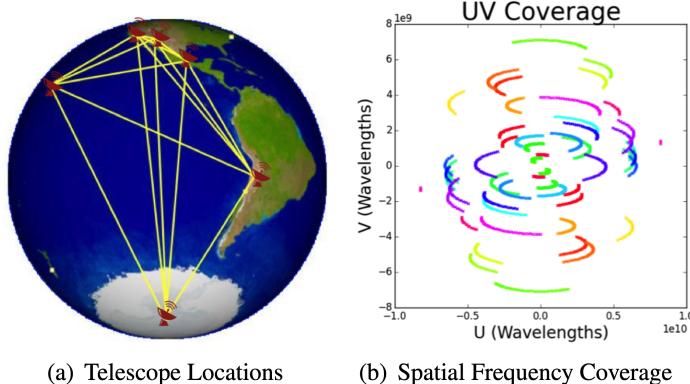


Figure 1: A sample of telescopes used in VLBI black hole imaging along with their corresponding baselines (a) and the frequency coverage of the telescope pairs (b) [6]. Sample points from the same telescope pair are marked in the same color. UV space coverage forms curves due to the baselines' rotation with Earth.

1.3 Black Hole Image Reconstruction Problem

The black hole image reconstruction problem can be formulated within the framework of interferometric imaging. In VLBI, the objective is to reconstruct an image of the black hole \mathbf{x} from

interferometric measurements \mathbf{y} . These measurements are connected to the true image through a **forward model** $f(\cdot)$ that depends on the arrangement of the telescopes. Mathematically, this relation is:

$$\mathbf{y} = f(\mathbf{x}) + \eta \quad (2)$$

where \mathbf{y} represents the visibilities, $f(\cdot)$ is the forward model that maps the image \mathbf{x} to the measurements based on the telescope configuration, and η denotes the noise in the measurements. The forward model $f(\cdot)$ typically involves a 2D Fourier transform, as visibilities are samples of the spatial frequency components of the image.

The problem of reconstructing an image \mathbf{x} from measurements \mathbf{y} with knowledge on the forward model $f(\cdot)$ is usually treated as a **Bayesian inference** problem [5]:

$$\mathbf{x}^* = \arg \max_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{y}) + \log p(\mathbf{x}), \quad (3)$$

where $p(\mathbf{x}|\mathbf{y})$ is the **data-fitting** term indicating how much the inferred solution matches the measurements, and $\log p(\mathbf{x})$ is the **prior** term on black hole images.

The primary challenge in this problem is that the UV plane coverage is incomplete and sparse, as shown in Fig. 1. This is caused by the limited number of telescopes, hence a limited number of available baselines on Earth. This information loss from sparse sampling renders the reconstruction a highly ill-posed problem, where there are infinitely many possible images that can well match the observed data. Therefore, additional constraints or prior information are needed to regularize the problem and guide the reconstruction towards a plausible solution. Moreover, understanding which part of the image might be incorrect when a precise solution is not possible is also crucial for scientific applications, which gives rise to the challenging problem of uncertainty quantification.

There are two primary types of priors used in this context: hand-crafted priors and data-driven priors. However, none of these priors works well in black hole image reconstruction.

Hand-crafted priors are manually designed based on known properties of images and physical phenomena. For example, total variation (TV) [7], assuming sparse gradients, is one of the most commonly used prior on natural images. These priors impose specific constraints to regularize the solution, making the reconstruction problem more tractable. While hand-crafted priors can provide significant improvements in image reconstruction, they require careful picking and hyperparameter tuning and may introduce human bias. Moreover, these priors may not fully capture the complex and unique structures found in black hole images, potentially limiting the accuracy of the reconstructions.

Data-driven priors, on the other hand, are learned from large datasets using advanced machine learning techniques. These priors can capture more complex and realistic patterns by leveraging the statistical properties of the data. Recent advancements in deep learning have led to the development of generative models that can learn such priors. The primary advantage of data-driven priors is their ability to adapt to the specific characteristics of the data, potentially leading to more accurate and detailed reconstructions. However, applying these priors to black hole imaging is challenging due to the scarcity of realistic black hole images. While simulated datasets can be used to train these models, there is a risk that the reconstructions might be biased towards the characteristics of the simulated data, which may not fully represent the true nature of black holes [8].

In conclusion, the major challenges in black hole image reconstruction are **uncertainty quantification** and **independence of explicit priors**.

2 Related Work

In this section, we will introduce previous work related to black hole image reconstruction. We will start with classic VLBI image reconstruction algorithms, CLEAN, and regularized maximum likelihood. Following the challenges in black hole imaging, we will then summarize recent advances in uncertainty quantification and image reconstruction without explicit priors respectively.

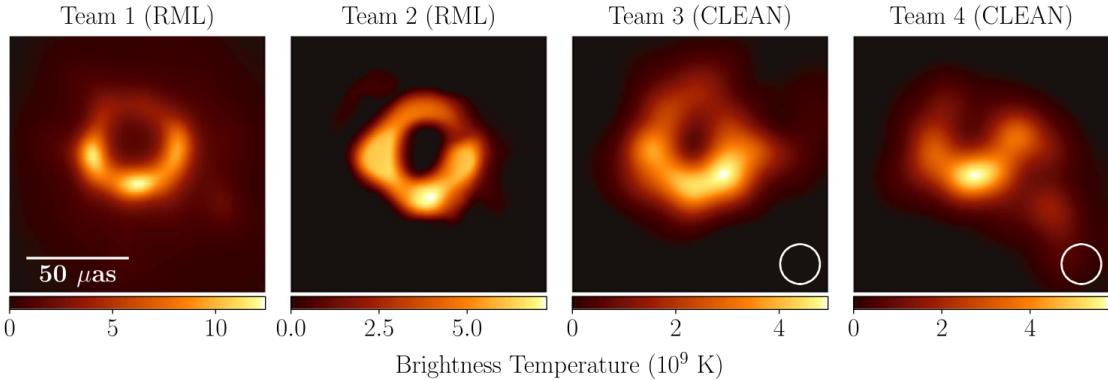


Figure 2: Images of M87 recovered by four different EHT teams using different methods [1]. The images share common low dimensional ring structure with enhanced brightness towards the south (lower part in the image).

2.1 Traditional Black Hole Image Reconstruction

To reconstruct black hole images from sparse and noisy VLBI measurements, astronomers developed classic methods that utilize hand-crafted priors to constrain the solution. Due to the lack of data-driven priors and reconstruction algorithms that are prior-free, these methods have been widely used in reconstructing the images of M87 and Sag A*, the first two black holes observed by the Event Horizon Telescope.

Classic imaging inverse methods in VLBI start with a 2D inverse Fourier transform on the observed visibilities (the sparse samples in UV space) and obtain a dirty image. The dirty image will then be deconvolved by the effects of the sparse 2D frequency coverage to yield a clean image. Following this strategy, the current standard method in VLBI image reconstruction is the **CLEAN** algorithm [9, 10, 11]. CLEAN treats the image as a set of independent point sources and iteratively solves for the location and flux of those sources. In every iteration, CLEAN will convolve the current set of point sources with an approximated point spread function of the imaging system (usually a Gaussian matched to the dirty image) and find the maximum difference between the dirty image and the convolved image. A new point source will be added at the location of the maximum difference image. CLEAN will repeat the following procedure until convergence. Eventually, the set of point sources obtained during this process will be the final output (the clean reconstructed image).

Another class of traditional VLBI image reconstruction is **regularized maximum likelihood (RML)**. RML methods are developed from the Bayesian probability perspective in (3) by maximizing the posterior probability of the reconstructed image given the measurements, the image forward model, and the image prior. It solves for the latent image by minimizing the following objective function [1],

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \alpha L(\mathbf{x}, f(\mathbf{y})) + \beta R(\mathbf{x}), \quad (4)$$

where $f(\cdot)$ is the image forward model, $L(\cdot)$ is the data fitting term calculating the difference between the measurements and the inferred measurements, $R(\cdot)$ is the prior on the image to be reconstructed. α, β are the hyperparameters weighing the effects of data-fitting and regularization. Large α will put too much weight on the data term, causing the image reconstructed to suffer from artifacts, while Large β will put too much weight on regularization, resulting in images not fitting the measurements well. The regularizers commonly used in RML include smoothness [6, 12], sparsity [13, 14, 15], total-variation (TV) [7], and entropy [16].

Traditional image reconstruction methods mentioned above can be susceptible to human bias introduced during data pre-processing, selection of regularizers, and hyperparameter tuning. To

address this challenge, the Event Horizon Telescope (EHT) collaboration employed a unique strategy. Four independent EHT teams were tasked with reconstructing the image of M87 from the April 2017 data using their own approaches. Four independent EHT teams were tasked with reconstructing the image of M87 from the April 2017 data using their own approaches. The resulting images, presented in Fig. 2, were obtained using different reconstruction algorithms and settings. Notably, all four reconstructions – two utilizing Regularized Maximum Likelihood (RML) and two primarily relying on the CLEAN algorithm – reveal a consistent ring structure with a brighter southern region (bottom of the image). This remarkable consistency suggests that the ring and the brighter south feature are intrinsic properties of M87, independent of the specific reconstruction method chosen by each team.

While averaging these independent reconstructions aimed to mitigate human bias, it's acknowledged that this approach can be labor-intensive. Additionally, even with averaging, a certain level of human bias might still be present in the final image due to its inherent existence within each team's initial reconstruction. Furthermore, quantifying the uncertainty in the reconstruction arising solely from human bias remains a challenge. These limitations highlight the urgent need for improved algorithms less susceptible to human influence and capable of quantifying uncertainty.

2.2 Uncertainty Quantification

Uncertainty quantification refers to the methods used to systematically and quantitatively determine the uncertainty in a model's output. In inverse problems, uncertainty quantification is crucial for assessing the confidence in solutions derived from observed measurements. It provides insights into how variations in input measurements and model assumptions affect the output, given that inverse problems are often sensitive to these factors and to measurement noise.

MCMC Sampling Markov Chain Monte Carlo (MCMC) is a commonly used technique for estimating the uncertainty in tasks related to image reconstruction. This method facilitates the estimation of a hidden image's posterior distribution through sampling techniques. Despite its advantages, MCMC has notable drawbacks, particularly in terms of computational speed. It becomes significantly slower when addressing high-dimensional inverse problems due to the increased data and model complexity, which elevate computational demands. Consequently, MCMC is often impractical for applications that require fast processing or are of a large scale.

Variational Bayesian Methods Variational inference is a prevalent technique employed to approximate intractable posterior distributions [17] [18]. Instead of directly computing the exact posterior distribution, variational Bayesian methods utilize a manageable set of density functions. They engage in solving an optimization problem to determine which function from this set best approximates the intended posterior distribution. Notably, these methods tend to be faster than traditional sampling techniques like MCMC, while frequently delivering performance that is on par. [19]. The effectiveness of variational methods hinges on the chosen family of density functions' capacity to model the complexity of the target distribution. A widely used method to streamline the modeling procedure is the mean-field approximation[20], which posits that the distributions of individual hidden variables are independent. Additionally, these density functions can be parameterized through neural networks, such as flow-based generative models, offering a versatile and potent means to encapsulate the intricacies of complex distributions.[21].

Bayesian Neural Networks Conventional deep learning models usually deliver just a singular estimate of the hidden image. On the other hand, Bayesian neural networks[22, 23], as discussed in Mackay (1995), offer a significant advancement by precisely quantifying the uncertainty involved in image reconstructions[24]. In these networks, the weights are treated as probability distributions, resulting in varied predictions with each execution of the network. This probabilistic modeling allows Bayesian neural networks to offer a more nuanced interpretation of image data by encapsulating the inherent uncertainty. While these networks often achieve remarkable results, their dependence on supervised learning confines their ability to represent uncertainty to only what is present within the training data. This constraint can affect the generalizability and reliability of the reconstructions

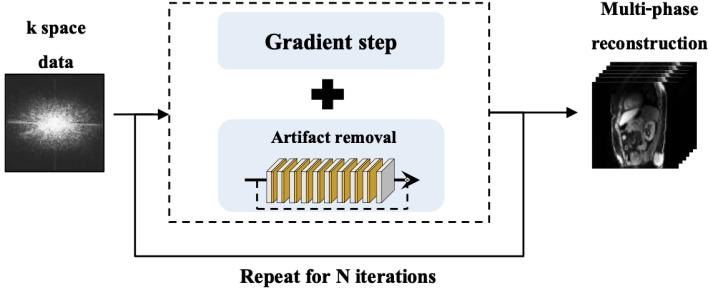


Figure 3: Regularization by artifact removal (RARE) [27]. RARE combines information from the CNN prior with that from the measurement operator to iteratively refine the solution.

in scenarios not well-represented in the training data.

Empirical Sampling An empirical alternative for generating multiple candidate reconstructions involves repeatedly solving a regularized inverse problem with varying regularizer hyper-parameters and different image initializations. This method was applied to characterize the uncertainty in the reconstructed image of the black hole M87* [1]. While the mean and standard deviation offer a preliminary assessment of uncertainty in our reconstructions, they don't necessarily represent a true posterior distribution based solely on the measurement data. This approach mainly captures the variability arising from choices made during reconstruction, like the selection of hyperparameters for the regularization technique. It doesn't account for uncertainties due to measurement noise or limitations caused by the inherently sparse data itself.

2.3 Image Reconstruction without Explicit Priors

Noise2Noise (N2N) Noise2Noise [25] trains deep neural networks for image restoration using only noisy image pairs, bypassing the need for clean target data. Mathematically, it exploits the insight that the expectation of a noisy distribution under certain noise models (e.g., zero-mean additive noise) converges to the clean signal. For training, the approach uses pairs of independently corrupted instances of the same image, assuming that the average of such noisy samples approximates the clean image. The method leverages a standard loss function, such as L2 (mean squared error), which in this setting, minimizes the expected value of the squared differences between the network's predictions and the noisy targets. This loss effectively encourages the network to predict the mean of the clean target distribution, as the noise, by assumption, cancels itself out on average. This enables the network to learn to denoise or reconstruct images without ever observing clean examples, simplifying data collection and extending applicability to situations where clean data are unavailable or costly to acquire. However, a significant limitation of N2N is its requirement for a large number of training images (typically on the order of thousands) to achieve good results. [26]. This poses a considerable challenge in cases where the available images do not come from the same distribution as the target images, where the available data consists of a few measurements from distinct underlying images, or when there is just a single measurement for each underlying image, lacking repeated observations. Variability in each measurement process can further complicate the task of reconstruction.

Regularization by Artifact Removal (RARE) RARE [27] introduces a groundbreaking framework in medical imaging that employs artifact-removal Convolutional Neural Networks (CNNs) trained on undersampled and noisy data, bypassing the need for the fully-sampled ground truth data that traditional methods typically require. This approach proves especially beneficial for applications such as dynamic MRI, where collecting extensive training data is often impractical. The framework utilizes an innovative Artifact2Artifact (A2A) training strategy, inspired by Noise2Noise, which

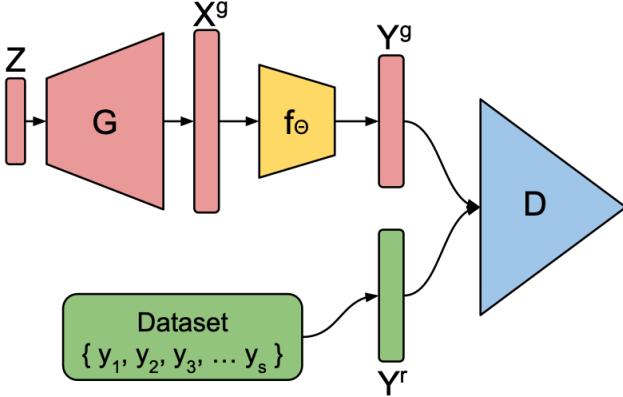


Figure 4: An illustration of AmbientGAN training [28]. The image generated by the generator G is passed through a simulated forward model to obtain a generated measurement Y^g . The task of the discriminator D is to judge whether the measurement is synthetic or realistic.

trains the model using pairs of undersampled, artifact-laden images directly from measurements. Extensive testing on both simulated and experimental MRI data has validated RARE’s ability to effectively harness learned priors and adapt to varied data conditions. A critical aspect of the framework is detailed in Fig. 3, which illustrates the structure of the Artifact Removal Network and the RARE approach. This figure highlights how RARE integrates the CNN prior with information from the measurement operator to iteratively refine the solution, offering a visual representation of the method’s operation.

One primary limitation of RARE is its reliance on the quality and diversity of the undersampled data used to train the artifact-removal networks. If the training data does not comprehensively represent the range of potential imaging scenarios, there is a risk that the model might not generalize well to new conditions. Additionally, the A2A strategy could be prone to residual errors if the artifact patterns in the training dataset do not fully encompass those encountered in actual diagnostic settings. Despite these challenges, RARE provides a robust framework for medical imaging in scenarios where traditional data acquisition is limited, marking a significant advancement for the field.

AmbientGAN The AmbientGAN approach addresses the challenge of training generative models with noisy or incomplete measurements, circumventing the need for fully-observed data samples. This method operates by introducing a simulated measurement process into the Generative Adversarial Network (GAN) training loop. Specifically, AmbientGAN modifies the training process of the discriminator. Rather than teaching it to identify the differences between real and generated images, it instructs the discriminator to discern between actual measurements and simulated measurements derived from generated images. This allows the model to learn from and adapt to the specific kinds of distortions present in the measurements. Despite its innovative approach, AmbientGAN requires a substantial number of measurement examples, typically around 10,000, to effectively train a high-quality generator. This necessity stems from the complexity and variability of the measurement distortions it must model to accurately reproduce the underlying data distribution. The training strategy of AmbientGAN is shown in Fig. 4.

Deep Image Prior (DIP) The main idea of the Deep Image Prior [29] is that the structure of a deep convolutional neural network (ConvNet) itself, even before any training on image datasets, is sufficient to capture essential low-level image statistics. This innate capability of Convolutional Networks (ConvNets) can be leveraged to tackle a range of image restoration challenges including denoising, super-resolution, and inpainting. It has been shown that even networks initialized randomly can act as effective priors for images. They employ this untrained network to effectively solve

inverse problems typically tackled with trained models. The method leverages the natural architectural biases of ConvNets—specifically their ability to model local coherence and handle different scales of image features—to reconstruct high-quality images from corrupted inputs. This approach challenges the conventional belief that the performance of deep learning in image restoration is solely due to learning from large datasets, suggesting instead that the network architecture itself plays a critical role.

However, DIP is susceptible to overfitting and necessitates the careful selection of a stopping criterion for the reconstruction process. This is generally manageable when example images are available to guide the determination of this stopping point. However, in scenarios where only noisy measurements are available, selecting an appropriate stopping condition can introduce significant human bias, potentially degrading the reconstruction quality.

3 Method

In this section, we will describe recent methods that address the challenges in black hole image reconstruction: (1) quantification of uncertainties and (2) independence to explicit priors. These methods (1) use a generative model in a variation Bayesian inference framework, (2) can perform uncertainty quantification by posterior sampling, and (3) do not rely on training data.

3.1 Deep Probabilistic Imaging

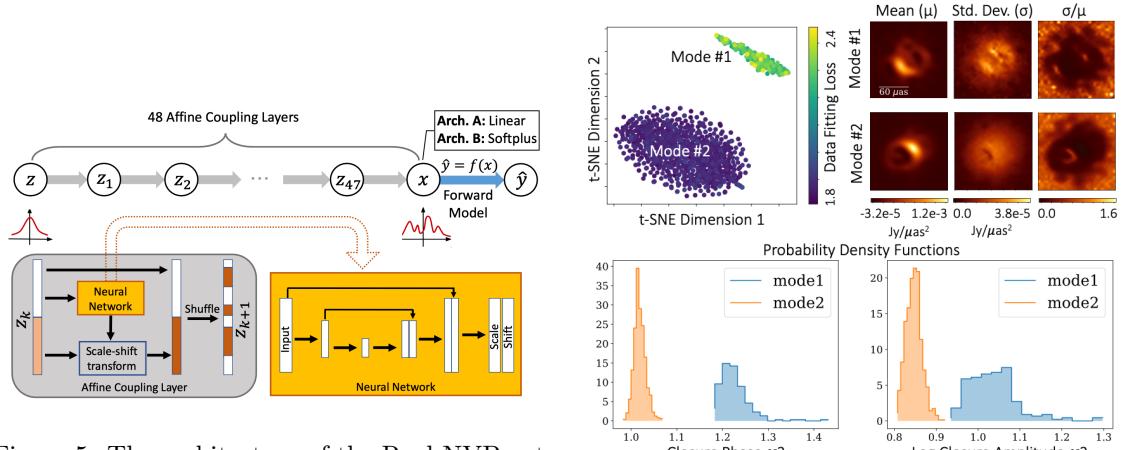


Figure 5: The architecture of the Real-NVP network used in DPI as the image generator [5]. This architecture is chosen to make the determinant calculation of the entropy term in (8) tractable.

Figure 6: DPI image reconstruction results with uncertainty map on synthetic data [5]. The images sampled from the inferred poster converged into two modes, which are 180 degrees rotational symmetric to each other.

To address the first challenge of quantifying uncertainty of the reconstructed black hole images, astronomers and computer scientists developed **Deep probabilistic Imaging (DPI)** [5], a variational approach that measures the posterior distribution in imaging inverse problems without training data. DPI utilizes an untrained image generation network to parameterize the posterior distribution of the latent image and optimizes the network weights using a loss function derived from the Kull-backLeibler (KL) divergence between the image posterior distribution and the model distribution.

To understand how this method works, we start from the parameterization of the posterior distribution using a flow-based generative network [30],

$$\mathbf{x} \sim q_\theta(\mathbf{x}) \Leftrightarrow \mathbf{x} = G_\theta(\mathbf{z}), \quad \mathbf{z} \sim N(0, \mathbf{I}), \quad (5)$$

where \mathbf{x} is the latent image, \mathbf{z} is a latent vector sample from a multi-variate standard normal distribution, and $G_\theta(\cdot)$ is the generative model, θ denotes the network weights. After the network is learned, such parameterization would enable effortless posterior sampling, from the uncertainty map can be easily calculated.

The goal of learning the network weights θ is to let the posterior parameterized by the network $q_\theta(\mathbf{x})$ be as close to the image posterior $p(\mathbf{x}|\mathbf{y})$ as possible. To measure the distance or similarity between two probabilistic distributions, we can use the Kull-backLeibler (KL) divergence. Thus the goal of the training is to minimize the KL divergence between $q_\theta(\mathbf{x})$ and $p(\mathbf{x}|\mathbf{y})$ [5]:

$$\begin{aligned} \theta^* &= \arg \min_{\theta} D_{\text{KL}}(q_\theta(\mathbf{x}) \| p(\mathbf{x}|\mathbf{y})) \\ &= \arg \min_{\theta} \int q_\theta(\mathbf{x}) [\log q_\theta(\mathbf{x}) - \log p(\mathbf{x}|\mathbf{y})] d\mathbf{x} \\ &= \arg \min_{\theta} \int q_\theta(\mathbf{x}) [\log q_\theta(\mathbf{x}) - \log p(\mathbf{y}|\mathbf{x}) - \log p(\mathbf{x})] d\mathbf{x} \\ &= \arg \min_{\theta} \mathbb{E}_{\mathbf{x} \sim q_\theta(\mathbf{x})} [-\log p(\mathbf{y}|\mathbf{x}) - \log p(\mathbf{x}) + q_\theta(\mathbf{x})]. \end{aligned} \quad (6)$$

Using the flow-based parameterization in (5), we can rewrite the training (6) in the following form [5]:

$$\theta^* = \arg \min_{\theta} \mathbb{E}_{\mathbf{z} \sim N(0, \mathbf{I})} \left[-\log p(\mathbf{y}|G_\theta(\mathbf{z})) - \log p(G_\theta(\mathbf{z})) + \pi(\mathbf{z}) - \log \left| \det \frac{dG_\theta(\mathbf{z})}{d\mathbf{z}} \right| \right]. \quad (7)$$

Approximating the expectation in (7) using a Monte Carlo approach and adding a weight parameter β to control the impact of the final entropy term, we obtain the following optimization problem [5]:

$$\theta^* = \arg \min_{\theta} \sum_{i=1}^N \left\{ L(\mathbf{y}, G_\theta(\mathbf{z}_i)) + \lambda R(G_\theta(\mathbf{z}_i)) - \beta \log \left| \det \frac{dG_\theta(\mathbf{z}_i)}{d\mathbf{z}_i} \right| \right\}, \quad (8)$$

where $\mathbf{z}_i \sim N(0, \mathbf{I})$ and β is the tuning parameter that controls the influence of the negative entropy term, which in turn controls the diversity of the generated samples.

The network is learned by drawing samples of latent vector \mathbf{z} and optimizing the objective function in (8). After the training converges, one can draw samples of \mathbf{z} and pass them through the learned network to get samples from the image posterior distribution $G_\theta(\mathbf{z})$. When the number of samples is large enough, the mean and standard deviation of the samples would be an accurate estimate of the reconstructed image and its uncertainty map.

In DPI, a Real-NVP [31] network is used as the image generator [5]. This architecture is selected to make the determinant calculation of the entropy term in (8) tractable. The The Jacobian matrix of Real-NVP is a multiplication of lower triangular matrices, which makes the determinant calculation very convenient.

Black hole images reconstructed using DPI are shown in Fig. 6. The posterior images converged into two modes in t-SNE latent space, which is a great example of multiple solutions fitting the measurements well in an ill-posed inverse problem. The mean images of the two modes are 180-degree rotational symmetric to each other, which matches the intrinsic physical constraints of VLBI. Mode 2 is the correct solution and has lower data-fitting loss, indicated by the color of the points. This synthetic experiment demonstrates DPI's power in accurately capturing the posterior distribution of the black hole image without any training data.

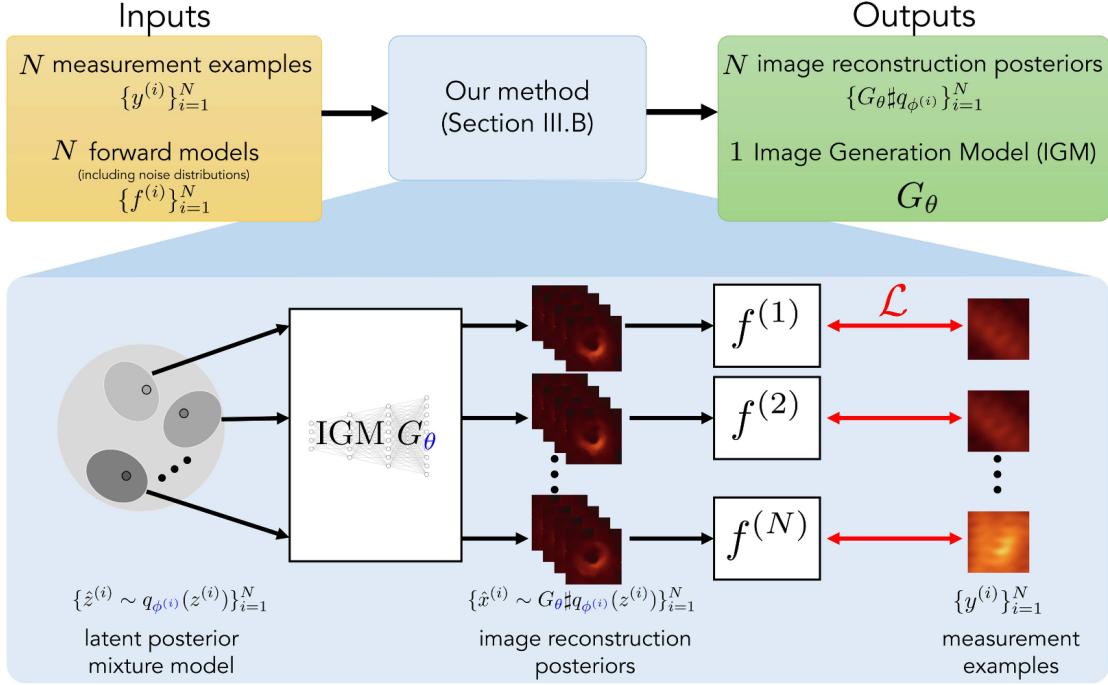


Figure 7: Overview of the image reconstruction method without explicit priors introduced in [33]. Given N measurements on the same black hole and N corresponding forward models, a shared image generation model (IGM) is learned as a regularization. N latent variable distributions are also optimized for posterior sampling.

3.2 Black Hole Image Reconstruction without Explicit Priors

Although DPI is able to solve an ill-posed inverse problem and quantify uncertainty, it still relies on hand-picked priors $R(\cdot)$ in the optimization (8). To solve this second challenge, computer scientists developed a new method that does not need on explicit priors based on the DPI framework [26, 32]. The basic assumption of these methods is different underlying images of the same black hole share the same low-dimensional structure, i.e. they are all ring-shaped and have similar size, but the location of bright spots along the ring might change with time. Following this assumption, one can learn a **image generation model (IGM)** jointly on multiple frames of the same black hole capture at different times. In this way, the IGM will be able to capture the shared low-dimensional structure of the black hole, which could be used as a task-specific regularizer in inverse problem-solving.

Similarly, an Evidence Lower Bound (ELBO) loss derived from KL divergence is also used to estimate how close the learned model distribution is to the image posterior distribution¹ [26]:

$$\text{ELBOPProxy}(G, q_\phi, \mathbf{y}) = \mathbb{E}_{\mathbf{z} \sim q_\phi(\mathbf{z})} [\log p(\mathbf{y}|G(\mathbf{z})) + \log P_Z(\mathbf{z}|G) - \log q_\phi(\mathbf{z})], \quad (9)$$

where, G is the IGM, $q_\phi(\cdot)$ is the distribution of latent vector, \mathbf{y} is the measurement, $P_Z(\cdot)$ is the prior distribution of the latent vector (chosen to be a standard normal distribution here for simplicity). After using Monte Carlo approximation and applying regularization on the IGM, we

¹Due to the length limit, we will skip the derivation of the ELBO loss in (9). The basic idea of utilizing KL divergence and flow-based network parameterization is the same as DPI. See [26] for detailed derivations.

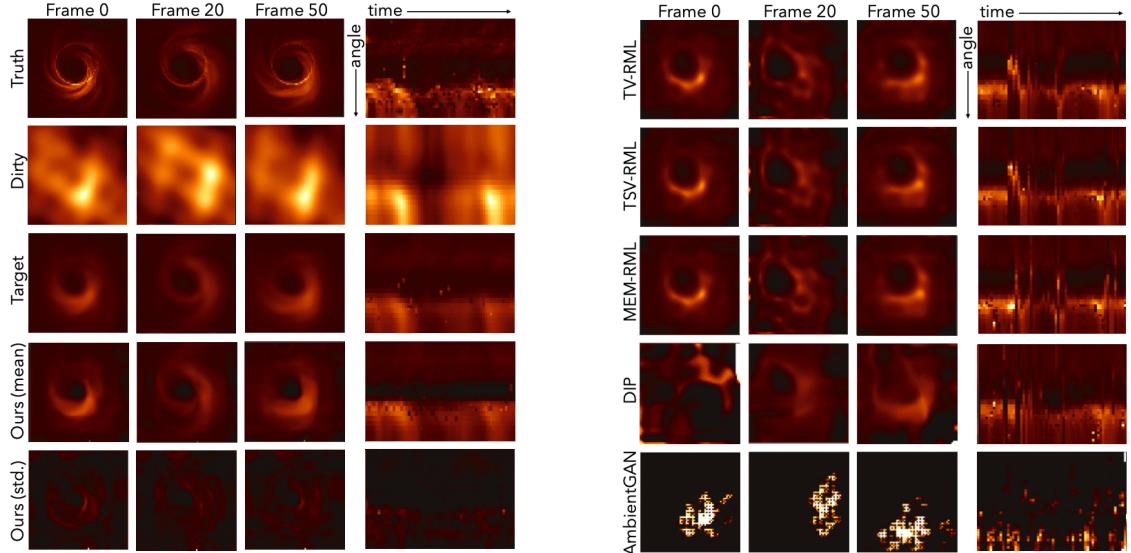


Figure 8: Reconstructions of M87* using the introduced method [34]. As the introduced method demonstrates success in both low and high frequency feature recovery and temporal consistency, while classic methods and modern methods all suffers from various types of artifacts and temporal inconsistency.

get the optimization problem of the IGM training [26]:

$$\{\theta^*, \phi_1^*, \dots, \phi_N^*\} = \arg \min_{\theta, \{\phi_i\}_{i=1}^N} \frac{1}{N} \sum_{i=1}^N \text{ELBOProxy}(G_\theta, q_{\phi_i}, \mathbf{y}_i) + \log p(G_\theta), \quad (10)$$

where θ is the weights of IGM, $\{\phi_i\}_{i=1}^N$ is the parameters of the latent vector distribution on the N measurements of the same black hole, and $\log p(G_\theta)$ is the regularization on the network weights, which is implemented in practice as sparse network weights and dropout during training.

Illustrated in Fig. 7, this method takes in N measurements of the same black hole captured at different times and learns the shared IGM G_θ and latent vector distribution parameters $\{\phi_i\}_{i=1}^N$ simultaneously. Posterior image samples can be efficiently obtained by leveraging the latent vector distributions. We can sample from these latent distributions and decode the samples using the IGM to generate posterior image samples. Similarly, applying a Monte Carlo approach allows us to compute the mean and standard deviation images.

Results of this method applied to synthetic VLBI black hole measurements are shown in Fig. 8. The discussed method can capture the low-frequency ring structure and part of the high-frequency spiral structures. In the right column, the discussed method can recover the temporal consistent trajectory of the bright spots. The baseline methods fail to recover a plausible ring structure and all suffer from temporal inconsistent trajectory. It is worth mentioning that the two modern methods, DIP and AmbientGAN, failed drastically. These results demonstrate the power of methods in [26, 32] to fully reconstruct black hole images by leveraging shared structures without explicit priors.

4 Limitations and Future Directions

Despite the success of the state-of-the-art methods discussed in this survey, there are still several limitations. In this section, we will discuss the limitations of these methods and provide future research directions accordingly.

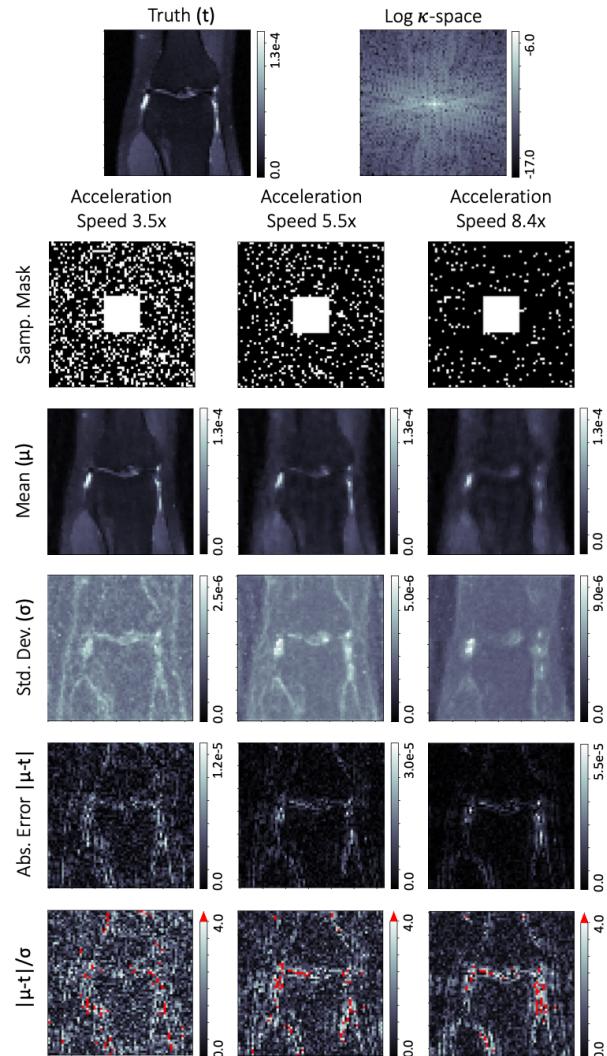


Figure 9: Reconstructed image of synthetic compressed sensing MRI with DPI [5]. The reconstructed image (Mean) agrees with the ground truth well and the uncertainty map (Std. Dev.) matches the absolute errors.

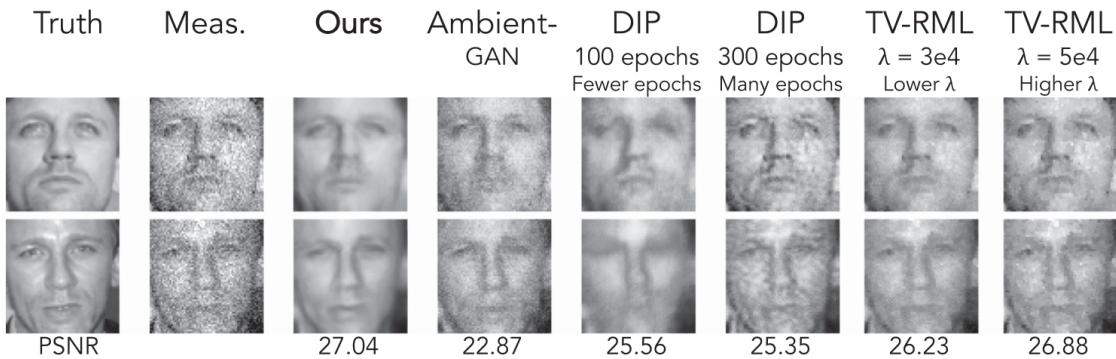


Figure 10: Comparison of the reconstruction method without explicit priors [26] with previous methods on image denoising. The current methods outperformed previous methods both quantitatively and qualitatively.

The first limitation of the methods discussed above is the interpretability of uncertainty. Although these methods are able to directly sample from the image posterior distribution and calculate the uncertainty map via a Monte Carlo approach, they cannot understand where the uncertainty is coming from. For instance, in computer vision, the two major components of uncertainty are **aleatoric uncertainty**, which models noise in the measurements, and **epistemic uncertainty**, which captures noise in the model [35]. Posterior sampling cannot directly divide uncertainty into those two classes therefore astronomers cannot find the biggest source of uncertainty to improve the model or the data collection process. Future methods should aim at understanding the source of uncertainty in the reconstruction image, this not only helps astronomers know how much they can trust the result in astrophysical studies but also provides a direct reference for computer scientists and engineers to better design the VLBI imaging system and reconstruction algorithm.

Another drawback of [5, 26, 32] is the lack of testing on realistic data. Although [5] did test the method on real EHT measurements of M87, it only tested on individual frames, while the others' results are still on simulation. The **next generation Event Horizon Telescope (ngEHT)** aims to study the dynamic properties of black holes by observing black holes over a time span of months. Reconstructing black hole videos over such a long time period is more challenging since there are much more nonidealities in real data and the observation condition might change over time, causing the measurements to be temporal inconsistent.

The optimizations in (8) and (10) are expensive and time-consuming. Although the weights are learned over a sample number of samples, the training requires $\sim 20,000$ epochs to converge. The parameters and network structure ablations might take more time to yield satisfactory results. This hinders scientists from quickly obtaining first-hand results and adjusting observation settings to take better black hole images. This might lead to the miss of optimum observation time for certain phenomena. Therefore, it is in great demand that faster reconstruction methods with uncertainty quantification be developed.

Looking at 2D black hole images is not the best way to study the dynamic properties of the black hole. Black holes are 3D systems with hot emissive gas orbiting around them. The ring visible in black hole images is the hot gas outside the black hole's event horizon. To study the dynamics of these emissive gases, it is desirable to directly perform **3D emission dynamics reconstruction**. Even though we can view black holes from a single point of view, this is possible because light rays are curved by the strong gravitational pull of the black hole, therefore we are able to collect light emitted behind the black hole. Inspired by Neural Radiance Fields (NeRF) [36], by adding gravitational lensed ray tracing into the VLBI forward model, one is able to reconstruct the 3D emission dynamics of the black hole over the observation period [37]. In 2023, astronomers performed a 3D reconstruction of an emission flare of Sag A*, the supermassive black hole at the center of the Milky Way, for the very first time [38]. If these methods could be combined with the variational framework with uncertainty quantification discussed in Section 3, it would provide a powerful tool to directly study the 3D emission dynamics around black holes.

The variational framework discussed in this survey can be generalized to any imaging inverse problems given knowledge of the forward model. Apart from black hole image reconstruction, there are image denoising, phase retrieval, compressed sensing MRI, X-ray CT, etc. These scientific imaging inverse problems all share the same challenge: loss of information due to physical constraints of the forward model. The methods discussed in this paper can also be applied to these ill-posed inverse problems because the scientific requirements for image reconstruction methods are the same in every scientific application: uncertainty quantification and independence of data-driven priors. For instance, DPI also shows promising performance on compressed sensing MRI image reconstruction [5], where the number of k-space (2D Fourier space) samples is reduced to save time. This is very similar to the information loss in VLBI UV space sampling. In Fig. 9, both the reconstructed image (Mean) and the uncertainty (Std. Dev.) match with the ground truth and absolute errors. In Fig. 10, [26] also outperforms previous methods on image denoising tasks, this could potentially be very useful for the enhancement of noisy medical images. For example, OCT images of the human retina suffer from high speckle noise, which hinders doctors from fast and accurate diagnosis.

5 Conclusions

In this survey, we reviewed and discussed various methods used in black hole imaging, encompassing both classic methods and modern learning-based approaches. We began with the scientific value and challenges of black hole imaging, introducing Very Long Baseline Interferometry (VLBI), the technique employed to produce measurements for reconstructing black hole images. The VLBI measurements are 2D frequency samples based on paired telescope data.

We then formulated black hole image reconstruction as a Bayesian inference problem and explained the failure of explicit priors. Hand-crafted priors introduce substantial human bias, while data-driven priors are infeasible due to data scarcity. The primary challenge is the ill-posedness of the inverse problem, arising from the sparse, non-convex, and noisy nature of VLBI measurements. We pointed out that in such a problem setting, uncertainty quantification and independence of explicit priors are the major challenges, and summarized previous work addressing those challenges.

We introduced two methods to address these challenges. The first method, Deep Probabilistic Imaging (DPI), utilizes a variational framework with a low-based generative model learned in an unsupervised manner. DPI estimates the posterior distribution of images that have not been directly observed by learning an image generator on a single measurement guided by a loss function encompassing data likelihood, an image prior, and negative entropy. DPI estimates the posterior distribution of the reconstructed image and is capable of identifying multiple potential solutions and quantifying uncertainty.

The second method solves ill-posed imaging inverse problems without established image priors or confirmed examples, by taking advantage of shared structure across multiple captures of the same object. This technique collectively solves various inverse problems by leveraging the interconnected structural elements of different images of the same black hole instead of depending on the spatial configuration of individual images. It employs a unified image generator learned with a loss function inspired by the Evidence Lower Bound (ELBO). This innovative approach, even when working with a limited dataset of flawed measurements, successfully tackles complex inverse challenges such as noise reduction, phase recovery, and the reconstruction of black hole imagery, underscoring the potential for the autonomous discovery of new scientific structures.

Our survey emphasized the crucial role of independence from data-driven biases and the quantification of uncertainties in advancing scientific studies. These innovations facilitate the creation of accurate black hole imagery, mitigating the impact of data scarcity and noise, and thus offer a strong solution to prevalent challenges in the field.

Lastly, we provide some potential future work in this field. As research progresses, it's crucial to assess whether current methods maintain their effectiveness when applied to realistic data collected over extended periods with the next-generation Event Horizon Telescope (ngEHT). There is also an opportunity to refine our understanding of the uncertainties these methods produce, identifying what originates from the data and what is influenced by the algorithms. Expanding the scope of our inquiry, incorporating state-of-the-art variational inference techniques and network architectures could potentially improve the quality and efficiency of image reconstruction. Furthermore, employing orbit dynamics and gravitational lensed ray tracing could enable the direct reconstruction of 3D emission dynamics around black holes, moving beyond traditional 2D imaging. This could provide more comprehensive insights into the behaviors and characteristics of astronomical phenomena.

A Contribution Statement

The contributions of each team member are listed below:

Tianao Li: Team leader, in charge of topic selection, Related Work (Sec. 2.1), Method (Sec. 3), Limitations and Future Directions (Sec. 4).

Yuxuan Li: Team member, in charge of Related Work (Sec. 2.2, 2.3) and Conclusions (Sec. 5).

Zhiyao Li: Team member, in charge of Background (Sec. 1).

References

- [1] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Bacsko, D. Ball, M. Baloković, J. Barrett, D. Bintley, *et al.*, “First m87 event horizon telescope results. iv. imaging the central supermassive black hole,” *The Astrophysical Journal Letters*, vol. 875, no. 1, p. L4, 2019.
- [2] E. H. TelescopeCollaboration, R. Azulay, A.-K. Bacsko, S. Britzen, G. Desvignes, R. Eatough, R. Karuppusamy, J.-Y. Kim, M. Kramer, T. Krichbaum, *et al.*, “First m87 event horizon telescope results. ii. array and instrumentation,” *apjl*, vol. 875, no. 1, p. L2, 2019.
- [3] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and synthesis in radio astronomy*. Springer Nature, 2017.
- [4] V. L. Fish, M. D. Johnson, R.-S. Lu, S. S. Doeleman, K. L. Bouman, D. Zoran, W. T. Freeman, D. Psaltis, R. Narayan, V. Pankratius, *et al.*, “Imaging an event horizon: mitigation of scattering toward sagittarius a,” *The Astrophysical Journal*, vol. 795, no. 2, p. 134, 2014.
- [5] H. Sun and K. L. Bouman, “Deep probabilistic imaging: Uncertainty quantification and multi-modal solution characterization for computational imaging,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 2628–2637, 2021.
- [6] K. L. Bouman, M. D. Johnson, D. Zoran, V. L. Fish, S. S. Doeleman, and W. T. Freeman, “Computational imaging for vlbi image reconstruction,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 913–922, 2016.
- [7] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [8] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, “Plug-and-play priors for model based reconstruction,” in *2013 IEEE global conference on signal and information processing*, pp. 945–948, IEEE, 2013.
- [9] J. Högbom, “Aperture synthesis with a non-regular distribution of interferometer baselines,” *Astronomy and Astrophysics Supplement*, Vol. 15, p. 417, vol. 15, p. 417, 1974.
- [10] B. Clark, “An efficient implementation of the algorithm ‘clean’,” *Astronomy and Astrophysics*, vol. 89, no. 3, Sept. 1980, p. 377, 378., vol. 89, p. 377, 1980.
- [11] F. Schwab, “Relaxing the isoplanatism assumption in self-calibration: applications to low-frequency radio interferometry,” *Astronomical Journal (ISSN 0004-6256)*, vol. 89, July 1984, p. 1076-1081., vol. 89, pp. 1076–1081, 1984.
- [12] A. A. Chael, M. D. Johnson, R. Narayan, S. S. Doeleman, J. F. Wardle, and K. L. Bouman, “High-resolution linear polarimetric imaging for the event horizon telescope,” *The Astrophysical Journal*, vol. 829, no. 1, p. 11, 2016.
- [13] Y. Wiaux, L. Jacques, G. Puy, A. M. Scaife, and P. Vandergheynst, “Compressed sensing imaging techniques for radio interferometry,” *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, pp. 1733–1742, 2009.
- [14] Y. Wiaux, G. Puy, Y. Boursier, and P. Vandergheynst, “Spread spectrum for imaging techniques in radio interferometry,” *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 2, pp. 1029–1038, 2009.
- [15] K. Akiyama, K. Kuramochi, S. Ikeda, V. L. Fish, F. Tazaki, M. Honma, S. S. Doeleman, A. E. Broderick, J. Dexter, M. Mościbrodzka, *et al.*, “Imaging the schwarzschild-radius-scale structure of m87 with the event horizon telescope using sparse modeling,” *The Astrophysical Journal*, vol. 838, no. 1, p. 1, 2017.

- [16] R. Narayan and R. Nityananda, “Maximum entropy image restoration in astronomy,” *Annual review of astronomy and astrophysics*, vol. 24, no. 1, pp. 127–170, 1986.
- [17] P. Arras, P. Frank, R. Leike, R. Westermann, and T. A. Enßlin, “Unified radio interferometric calibration and imaging with joint uncertainty quantification,” *Astronomy & Astrophysics*, vol. 627, p. A134, 2019.
- [18] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, “Variational inference: A review for statisticians,” *Journal of the American statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.
- [19] S. Gershman, M. Hoffman, and D. Blei, “Nonparametric variational inference,” *arXiv preprint arXiv:1206.4665*, 2012.
- [20] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, “An introduction to variational methods for graphical models,” *Machine learning*, vol. 37, pp. 183–233, 1999.
- [21] D. Rezende and S. Mohamed, “Variational inference with normalizing flows,” in *International conference on machine learning*, pp. 1530–1538, PMLR, 2015.
- [22] Y. Gal *et al.*, “Uncertainty in deep learning,” 2016.
- [23] D. J. MacKay, “Bayesian neural networks and density networks,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 354, no. 1, pp. 73–80, 1995.
- [24] Y. Xue, S. Cheng, Y. Li, and L. Tian, “Reliable deep-learning-based phase imaging with uncertainty quantification,” *Optica*, vol. 6, no. 5, pp. 618–629, 2019.
- [25] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, “Noise2noise: Learning image restoration without clean data,” *arXiv preprint arXiv:1803.04189*, 2018.
- [26] O. Leong, A. F. Gao, H. Sun, and K. L. Bouman, “Discovering structure from corruption for unsupervised image reconstruction,” *IEEE Transactions on Computational Imaging*, vol. 9, pp. 992–1005, 2023.
- [27] J. Liu, Y. Sun, C. Eldeniz, W. Gan, H. An, and U. S. Kamilov, “Rare: Image reconstruction using deep priors learned without groundtruth,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1088–1099, 2020.
- [28] A. Bora, E. Price, and A. G. Dimakis, “Ambientgan: Generative models from lossy measurements,” in *International conference on learning representations*, 2018.
- [29] D. Ulyanov, A. Vedaldi, and V. Lempitsky, “Deep image prior,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 9446–9454, 2018.
- [30] D. Rezende and S. Mohamed, “Variational inference with normalizing flows,” in *Proceedings of the 32nd International Conference on Machine Learning* (F. Bach and D. Blei, eds.), vol. 37 of *Proceedings of Machine Learning Research*, (Lille, France), pp. 1530–1538, PMLR, 07–09 Jul 2015.
- [31] L. Dinh, J. Sohl-Dickstein, and S. Bengio, “Density estimation using real NVP,” in *International Conference on Learning Representations*, 2017.
- [32] A. F. Gao, O. Leong, H. Sun, and K. L. Bouman, “Image reconstruction without explicit priors,” in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, IEEE, 2023.

- [33] O. Leong, A. F. Gao, H. Sun, and K. L. Bouman, “Discovering structure from corruption for unsupervised image reconstruction,” *IEEE Transactions on Computational Imaging*, vol. 9, pp. 992–1005, 2023.
- [34] Y. Y. Lin, A. F. Gao, and K. L. Bouman, “Imaging an evolving black hole by leveraging shared structure,” in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2475–2479, 2024.
- [35] A. Kendall and Y. Gal, “What uncertainties do we need in bayesian deep learning for computer vision?,” *Advances in neural information processing systems*, vol. 30, 2017.
- [36] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, “Nerf: Representing scenes as neural radiance fields for view synthesis,” *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.
- [37] A. Levis, P. P. Srinivasan, A. A. Chael, R. Ng, and K. L. Bouman, “Gravitationally lensed black hole emission tomography,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 19841–19850, 2022.
- [38] A. Levis, A. A. Chael, K. L. Bouman, M. Wielgus, and P. P. Srinivasan, “Orbital polarimetric tomography of a flare near the sagittarius a* supermassive black hole,” *Nature Astronomy*, pp. 1–9, 2024.