Article

Modeling the Complexity of Elliptic Black Hole Solution In 4D Using Hamiltonian Monte Carlo with Stacked Neural Networks

Ehsan Hatefi [*] , Armin Hatefi , Roberto Javier Lopez-Sastre

Article

Modeling the Complexity of Elliptic Black Hole Solution In 4D Using Hamiltonian Monte Carlo with Stacked Neural Networks

Armin Hatefi [1], Ehsan Hatefi [2]* and Roberto J. López-Sastre [2]

[1]   Department of Mathematics and Statistics, Memorial University of Newfoundland, St John's, NL, Canada; ahatefi@mun.ca

[2]   University of Alcalá, Department of Signal Theory and Communications, Research group GRAM, Alcalá de Henares, Spain; robertoj.lopez@uah.es

*   Correspondence: ehsan.hatefi@uah.es or ehsanhatefi@gmail.com

Abstract: In this paper, we study the black hole solution of self-similar gravitational collapse in the Einstein-axion-dilaton system for the elliptic class in four dimensions. The solution is invariant under space-time dilation, which is combined with internal SL(2,R) transformations. Due to the complex and highly nonlinear pattern of the equations of motion in the physics of black holes, researchers typically have to use various numerical techniques to make the equations tractable to estimate the parameters and the critical solutions. To this end, they have to ignore the numerical measurement errors in estimating the parameters. To our knowledge, for the first time in the literature on axion-dilation systems, we propose to estimate the critical collapse functions in a Bayesian framework. We develop a novel methodology to translate the modelling of the complexity of the elliptic black hole to a sampling problem using Hamiltonian Monte Carlo with stacked neural networks. Unlike methods in the literature, this probabilistic approach enables us not only to recover the available deterministic solution but also to explore possibly all physically distinguishable self-similar solutions that may occur due to numerical measurement errors.

Keywords: mathematical physics; black holes; statistical analysis; computational physics

1. Introduction

Black holes are known by their mass, angular momentum and their charge. Choptuik [1] illustrates that there may be another parameter explaining the gravitational collapse solutions. Christodolou [2] elaborated on the spherically symmetric collapse of the scalar field. Choptuik [1] empirically showed the critical behaviour of the solutions that illustrated the discrete self-similarity for the real scalar field gravitational collapse. The gravitational collapse solution shows space-time self-similarity so that dilations occur. The critical solution provides the scaling law. We can illustrate the initial condition of the real scalar field by parameter p, which is related to the field amplitude. Let $p = p_{\text{crit}}$ define the critical solution. The black hole is then formed when p gets bigger than p_{crit}. If $p > p_{\text{crit}}$, then one finds out the mass of the black hole or the Schwarzschild radius by a scaling law as follows

$$r_S(p) \propto M_{\text{bh}}(p) \propto (p - p_{\text{crit}})^{\gamma}. \tag{1}$$

The critical exponent for a real scalar field in four dimensions is found to be $\gamma \simeq 0.37$ [1,3,4], while for the other dimensions ($d \geq 4$), the mass of black hole is given by [5,6] as

$$r_S(p) \propto (p - p_{\text{crit}})^{\gamma}, \quad M_{\text{bh}}(p) \sim (p - p_{\text{crit}})^{(D-3)\gamma}. \tag{2}$$

The numerical optimization for different matter content can be read from [7–12]. The collapse solutions of the perfect fluid had been studied in [5,13–15] and its critical exponent $\gamma \simeq 0.36$ was also found in [14]. It has also been discussed in [16] that γ may have a universal value for all the fields that are coupled to gravity in four dimensions. One method for obtaining the critical exponent is based on perturbations of self-similar solutions. This method was established in [5,15,17]. Let h be any field, its perturbation is given by the following

$$h = h_0 + \varepsilon\, h_{-\kappa}, \tag{3}$$

where $h_{-\kappa}$ has the scaling $-\kappa \in \mathbb{C}$ which is related to different modes. The most relevant mode κ^* is related to the biggest value of $\mathrm{Re}(\kappa)$. The minus sign demonstrates a growing mode near the black hole formation time $t \to 0$. It was proved in [5,15,17] that κ^* can be related to Choptuik exponent by the following relation

$$\gamma = \frac{1}{\mathrm{Re}\,\kappa^*}. \tag{4}$$

The axial symmetry was studied in [18]. On the other hand, Alvarez-Gaume et al. [19] explore other solutions as well as shock waves. The critical exponent $\gamma \simeq 0.2641$ for the axion-dilaton system in four dimensions was first discovered in [20]. The analysis of [3,21] had been done in four dimensions for the elliptic case; however, their techniques can be used in hyperbolic and parabolic cases as well as in other dimensions, for further results see [22,23].

As the first motivation for the critical collapse of the axion-dilaton system, one can consider the AdS/CFT correspondence [24–27], relating the Choptuik exponent, the imaginary part of quasinormal modes and also the dual conformal field theory [28]. In the context of AdS/CFT, it is desirable to consider the spaces which tend asymptotically to $AdS_5 \times S^5$. From the bosonic fields of the theory, one can employ the system of axion-dilaton and the self-dual 5-form to analyze the critical collapse solutions in different dimensions. The other motivation of the axion-dilaton system is actually related to the holographic description of black hole formation [13,29]. The next motivation could potentially be the application of the system to black hole physics [30,31] as well as the important role of S-duality in self similar solutions [32].

The families of continuous self-similar solutions were explored by [33] in four and five dimensions for all classes of SL(2,R). These solutions are the extended results of [34,35]. More specifically, [36] studied the perturbations and reproduced the existing value [20] of $\gamma \sim 0.2641$ in four dimensions. Some other critical exponents for four and five dimensions were proved in [23]. Hatefi and Hatefi [37] proposed the Fourier-based regression models for the nonlinear critical collapse functions. To address the challenges of [37], Hatefi and Hatefi [38] employed truncated power basis, natural spline and penalized B-spline regression models to estimate the non-linear critical collapse functions. Recently, Hatefi et al. [39] applied artificial neural networks to analyze the black hole solutions in parabolic class in higher dimensions in detail.

In this paper, we propose a new methodology, which allows us to approach the problem of the complexity of elliptical black holes in 4 dimensions by means of Bayesian statistical modeling. Technically, we address the problem of modeling the self-similar solutions for the described spherical gravitational collapse [3,21] in four dimensions, using Hamiltonian Monte Carlo with Stacked Neural Networks. The modeling of black holes in the 4-dimensional elliptic case is reduced to a set of differential equations, which have always been approached by numerical methods that try to find a unique solution that satisfies the initial conditions. On the contrary, we argue that it is possible to attack the problem with Bayesian statistical modeling of the differential equations. In order to simplify the equations and the parameters of the equations of motion, researchers often have to use a variety of numerical approaches due to the complicated and highly nonlinear form of the equations of motion in the physics of black holes (e.g. [22,38,39]). For this reason, while calculating the parameters of the equations of motion, they must disregard the measurement errors imposed by the numerical calculations.

To the best of our knowledge, for the first time in the literature on the axion-dilation black hole, we investigate the equations of motion of the critical collapse functions within a Bayesian framework. This Bayesian framework enables us to address the measurement errors imposed by the numerical calculations in the estimation process. The parameters of the equations of motion are handled as random variables due to this statistical modelling. Because of this, we can derive the posterior distributions for these parameters, which we can then utilize as legitimate priors for a collection of stacked neural network-based solvers that will calculate the critical collapse function. In a nutshell, this paper describes a novel pipeline that starts solving the elliptic black hole equations in 4 dimensions using Hamiltonian Monte Carlo statistical modelling. Then, the posterior distributions discovered, for the parameters of the equations, are used with stacked neural networks to estimate the critical collapse functions. Unlike methods in the literature, this probabilistic perspective enables us to recover the deterministic solution in the literature and find all possible solutions that may occur due to numerical measurement errors in the parameter estimation phase. This paper paves the path to view the problem from a stochastic rather than a deterministic perspective.

This paper is organized as follows. Section 2 discusses the effective action of the axion-dilaton system in four dimensions, its equations of motion, and the initial conditions for the axion-dilaton system that come from the requirement of continuous self-similarity. Section 3 describes our novel approach for modelling the complexity of elliptic black hole solution in 4d from Hamiltonian Monte Carlo with stacked neural networks. Through extensive numerical studies, we develop the Bayesian posterior mean and median neural networks and the 95% credible intervals for all the critical functions in the elliptic class 4d in Section 4. Finally, in Section 5, we present the summary and concluding remarks.

2. The Relevant System

The axion-dilaton fields are combined into a single complex scalar field $\tau \equiv a + ie^{-\phi}$. The four-dimensional axion-dilaton (a, Φ) action is given by

$$S = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left(R - \frac{1}{2} \frac{\partial_a \tau \partial^a \bar{\tau}}{(\mathrm{Im}\,\tau)^2} \right), \tag{5}$$

where R is the scalar curvature. By taking variations from the metric and τ, one finds all the equations of motion as follows

$$R_{ab} = \frac{1}{4(\mathrm{Im}\,\tau)^2} \left(\partial_a \tau \partial_b \bar{\tau} + \partial_a \bar{\tau} \partial_b \tau \right), \tag{6}$$

$$\nabla^a \nabla_a \tau + \frac{i \nabla^a \tau \nabla_a \tau}{\mathrm{Im}\,\tau} = 0. \tag{7}$$

Note that the theory is classically invariant under SL(2,R) transformations, that is, if τ gets exchanged by

$$\tau \to \frac{a\tau + b}{c\tau + d}, \tag{8}$$

where $(a, b, c, d) \in R$, $ad - bc = 1$ and g_{ab} as well as the action remain to be invariant. This group can be broken to SL(2,Z) as a consequence of duality in String Theory (for review of duality see for instance [40–43]).

The spherically symmetric metric is given by [21] as

$$ds^2 = (1 + u(t, r)) \left(-b(t, r)^2 dt^2 + dr^2 \right) + r^2 d\Omega_{d-2}^2. \tag{9}$$

One can implement time re-definitions for (9), hence as in [21] we set $b(t, 0) = 1$ and regularity condition implies $u(t, 0) = 0$. The continuous self-similarity means the existence of a killing vector ξ that produces global scale transformation as follows

$$\mathcal{L}_\xi g_{ab} = 2 g_{ab}. \tag{10}$$

In spherical coordinates, one chooses $\xi = t\,\partial/\partial t + r\,\partial/\partial r$. Now by defining the scale invariant variable $z = -r/t$, self-similarity of the metric reflects the fact that all the functions $u(t,r), b(t,r)$ can be expressed in terms of z.

The action in (5) is SL(2,R)-invariant, so one compensates a scale transformation of (t,r) by an SL(2,R) transformation. Hence, by considering the change of variables to (t,z), one derives a differential condition (as originally pointed out in [34]) for $\tau(t,z)$ as

$$t\frac{\partial}{\partial t}\tau(t,z) = a + b\tau + c\tau^2 \,, \tag{11}$$

where all a, b, c are real numbers. The above equation has two roots which can be either two complex conjugate numbers, two real numbers or a double real root where they are related to compensating the scaling transformation in SL(2,R). In the elliptic class, the general form of the ansatz is given by

$$\tau(t,r) = i\frac{1 - (-t)^{i\omega}f(z)}{1 + (-t)^{i\omega}f(z)} \,, \tag{12}$$

where under a scaling transformation $t \to \lambda\,t$, $\tau(t,r)$ changes by a SL(2,R) rotation, which means that all equations are invariant under the following transformation

$$f(z) \to e^{i\theta}f(z)\,, \tag{13}$$

where $f(z)$ is a complex function that satisfies $|f(z)| < 1$ to be determined by solutions to non-linear ordinary differential equations, and ω is a real constant.

In the following, we describe briefly the derivation of the equations of motion for the elliptic class in four dimensions. Applying continuous self-similarity ansätze (12) to all the equations of motion (6) and (7), one finds the ordinary differential equations for $u(z), b(z), f(z)$. By taking the spherical symmetry, $u(z)$ can be expressed in terms of $b(z)$ and $b'(z)$ as

$$u(z) = \frac{z\,b'(z)}{(d-3)\,b(z)} \,. \tag{14}$$

Note that the first derivative of $u(z)$ can be eliminated from all equations of motion by the following constraint

$$\frac{qu'(z)}{2(1+u(z))} = \frac{4zf'(z)\bar{f}'(z) + 2i(w\bar{f}(z)f'(z) - wf(z)\bar{f}'(z))}{2(f(z)\bar{f}(z) - 1)^2} \,. \tag{15}$$

Hence, the equations of motion in the elliptic class in four dimensions are then given by

$$0 = b' + \frac{z(b^2 - z^2)}{b(-1 + |f|^2)^2}f'\bar{f}' - \frac{i\omega(b^2 - z^2)}{b(-1 + |f|^2)^2}(f\bar{f}' - \bar{f}f') - \frac{\omega^2 z|f|^2}{b(-1 + |f|^2)^2}, \tag{16}$$

$$
\begin{aligned}
0 = {} & f'' - \frac{z(b^2 + z^2)}{b^2(-1 + |f|^2)^2}f'^2\bar{f}' + \frac{2}{(1 - |f|^2)}\left(1 - \frac{i\omega(b^2 + z^2)}{2b^2(1 - |f|^2)}\right)\bar{f}f'^2 \\
& + \frac{i\omega(b^2 + 2z^2)}{b^2(-1 + |f|^2)^2}ff'\bar{f}' + \frac{2}{z}\left(1 + \frac{i\omega z^2(1 + |f|^2)}{(b^2 - z^2)(1 - |f|^2)}\right. \\
& \left. + \frac{\omega^2 z^4|f|^2}{b^2(b^2 - z^2)(1 - |f|^2)^2}\right)f' + \frac{\omega^2 z}{b^2(-1 + |f|^2)^2}f^2\bar{f}' + \\
& \frac{2i\omega}{(b^2 - z^2)}\left(\frac{1}{2} - \frac{i\omega(1 + |f|^2)}{2(1 - |f|^2)} - \frac{\omega^2 z^2|f|^2}{2b^2(-1 + |f|^2)^2}\right)f. \tag{17}
\end{aligned}
$$

The first equation of motion includes $b(z), f(z), f'(z)$ where b is a first-order linear in-homogeneous equation where its initial condition is given by $b(0) = 1$. The initial conditions for $f(z), f'(z)$ are known by applying the smoothness of the critical solution. Applying polar coordinate representation $f(z) = f_m(z)e^{if_a(z)}$, one can easily show that all equations are invariant under a global redefinition of the phase of $f(z)$. This implies that $f_a(0) = 0$. From the regularity condition at $z = 0$ and the residual symmetries in the equations of motions (17), one can derive the initial boundary conditions of the equations of motion by

$$b(0) = 1, f_m(0) = x_0, \qquad f'_m(0) = f'_a(0) = f_a(0) = 0 \tag{18}$$

3. Statistical Methods

Let $\mathbf{x}(t) = (x_1(t), \ldots, x_H(t))$ denote the solutions to a system of H differential equations (DEs), where $x_i(t)$ represents the solution to the i-th equation; henceforth, $x_i(t)$ is called the DE variable. Combining all the DE variables, we can model the system by

$$\frac{d}{dt}x_i(t) = g_i(\mathbf{x}(t)|`) , \tag{19}$$

where t denotes the space-time argument, ` represents the vector of all unknown parameters in the DE system with $x_i(0) = x_{i0}$ being the initial condition for the i-th DE variable.

3.1. The Likelihood Function

According to the complex and nonlinear pattern of the equations of motion, the true trajectory of the DE variables is not observable. Instead, one obtains the trajectory of the DE variables with measurement error implied by numerical experiments. To take this uncertainty into our statistical models, let $\mathbf{y}_i = (y_{i1}, \ldots, y_{in_i})$ denote the observed trajectory of DE variable $x_i(t)$ at n_i points where y_{ij} represents the observed value of $x_i(t)$ with measurement error at space-time point $j, j = 1, \ldots, n_i$. To take into account the errors involved in measuring the DE variables, we assume that y_{ij} follows Gaussian distribution with mean $x_i(t_j|`)$ and variance σ_i^2 for $i = 1, \ldots, H$; that is

$$y_{ij} \sim N\left(x_i(t_j|`), \sigma_i^2\right), \, j = 1, \ldots, n_i,$$

where $x_i(t_j|`)$ denotes the true value of the DE variable given the unknown parameters ` and t_j denotes the space-time point where we observed y_{ij} as the j-th value of the DE variable $x_i(t_j|`)$. Let $\mathbf{\Omega} = (`, \boldsymbol{\text{œ}})$ be the set of all unknown parameters of the Gaussian process, where $\boldsymbol{\text{œ}} = (\sigma_1, \ldots, \sigma_H)$. Given $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_H)$, the likelihood function of $\mathbf{\Omega}$ is given by

$$L(\mathbf{\Omega}|\mathbf{y}) = \prod_{i=1}^{H}\prod_{j=1}^{n_i}\left(\frac{1}{\sigma_i^2}\right)^{-1/2}\exp\left\{-\frac{(y_{ij} - x_i(t_j|`))^2}{2\sigma_i^2}\right\}. \tag{20}$$

The likelihood function (20) translates the DE system into a Gaussian process. Thus, one can estimate the unknown parameter $\mathbf{\Omega}$ by maximizing the likelihood function (20) and obtain the maximum likelihood estimates of $\mathbf{\Omega}$. According to the complex and nonlinear forms of the equations of motion, the likelihood function surface may lead to multiple isolated optimum points. Therefore, the likelihood function may appear very sensitive regarding parameter changes.

3.2. Hamiltonian Monte Carlo

We develop a Bayesian framework to estimate the unknown parameters of DE system (19). One must incorporate prior distribution for the model parameters $\mathbf{\Omega}$ in the Bayesian framework. Suppose $\pi(\mathbf{\Omega}; \mathbf{ff})$

denotes the prior distribution for $\mathbf{\Omega}$, where $\mathbf{ff}$ denotes the vector of all hyper-parameters of the prior distribution. From (20) and the prior distribution $\pi_0(\mathbf{\Omega}; \mathbf{ff})$, the posterior distribution $\mathbf{\Omega}$ is given by

$$\pi(\mathbf{\Omega}|\mathbf{y}) = \frac{L(\mathbf{\Omega}|\mathbf{y})\pi_0(\mathbf{\Omega}; \mathbf{ff})}{\int_{\mathbf{\Omega}} d\mathbf{\Omega} L(\mathbf{\Omega}|\mathbf{y})\pi_0(\mathbf{\Omega}; \mathbf{ff})}.$$

Due to the complex uncertainty involved in the multidimensional integrals of the marginal distributions, there is no closed form for the posterior distribution.

Markov Chain Monte Carlo (MCMC) is a well-established method in the Bayesian framework to make inferences about the posterior distribution of differential equations [44]. The MCMC approach translates the estimation problem into the sampling from the target non-normalized posterior distribution $\pi(\mathbf{\Omega}|\mathbf{y}) \propto L(\mathbf{\Omega}|\mathbf{y})\pi_0(\mathbf{\Omega}; \mathbf{ff})$. There is a rich literature on the properties of the MCMC methods [45,46]. MCMC employs various stochastic steps to construct a Markov chain from the target posterior distribution numerically. If the chain is run long enough, it is guaranteed that the MCMC converges to the true posterior distribution. It reaches a stationary state where the sampler enables us to generate samples from the posterior distribution.

Metropolis-Hastings (MH) is one of the most common MCMC approaches [45,47]. The MH uses a transition distribution $q(\mathbf{\Omega}^*|\mathbf{\Omega})$ computing the probability of making transition from $\mathbf{\Omega}$ to $\mathbf{\Omega}^*$ in the domain of the parameter space. MH then employs a stochastic step to accept or reject the transition. Let $\mathbf{\Omega}^{(t)}$ denote the t-th state of the MCMC. The probability of accepting the transition to candidate state $\mathbf{\Omega}^*$ is then calculated by

$$\min\left\{1, \frac{\pi(\mathbf{\Omega}^*|\mathbf{y})q(\mathbf{\Omega}|\mathbf{\Omega}^*)}{\pi(\mathbf{\Omega}|\mathbf{y})q(\mathbf{\Omega}^*|\mathbf{\Omega})}\right\}. \tag{21}$$

Random Walk (RW) is a common MH approach using a guess-and-check strategy to generate proposals from $\pi(\mathbf{\Omega}|\mathbf{y})$ in a neighbourhood of the current state of the MCMC. The RW is a popular approach because it is simple and intuitive to generate samples from a neighbourhood of the current state of the chain. Despite these advantages, the RW may appear very inefficient in high dimensions and complex systems where the posterior distribution may have multiple isolated modes. If the RW is tuned to take high random jumps, the MCMC chain may randomly jump from one low-density area in the parameter space to another. Consequently, this leads to a high probability of rejection and thus results in a biased estimate. On the other side, if RW is restricted to small jumps, the MCMC keeps accepting all the proposals and will be unable to visit other high-density posterior areas. The MCMC again leads to a biased estimate.

Hamiltonian Monte Carlo (HMC) can remedy the above challenges and find the posterior distribution in the DE systems [48,49]. The HMC augments the MH sampler with a momentum variable $\mathbf{\OE}$ to make the transition much more rapidly across the parameter space. This enables the MCMC to sample from almost all high-density areas and their neighbourhoods. Consequently, the MCMC estimates more effectively the posterior distribution. The HMC, as an iterative technique, consists of three steps in each iteration. Let $(\mathbf{\Omega}^{(t-1)}, \mathbf{\OE}^{(t-1)})$ be the current state of the MCMC in the t-th iteration. Thus, the (t)-th iteration of the HMC is carried out as follows:

1. **Momentum Initialization Step:** We initialize the candidate $\mathbf{\OE}^*$ from the prior distribution $p(\mathbf{\OE})$.
2. **Leapfrog Step:** This is the main step of the HMC consisting of three steps to upload simultaneously the DE parameters $\mathbf{\Omega}$ and momentum parameters $\mathbf{\OE}$. The leapfrog step is then alternated for L iterations where each iteration involves the (i)-(iii) steps:

(i) We update the half-step of the momentum parameter Œ using the gradient of the log-posterior distribution $\pi(\mathbf{\Omega}|\mathbf{y})$ scaled by tuning factor ϵ as

$$\text{Œ}^* = \text{Œ} + \frac{1}{2}\epsilon\frac{d}{d\mathbf{\Omega}}\log\pi(\mathbf{\Omega}|\mathbf{y}), \tag{22}$$

(ii) The half-step update (22) is used to full-step update of the DE parameters $\mathbf{\Omega}$ by

$$\mathbf{\Omega}^* = \mathbf{\Omega} + \epsilon M^{-1}\text{Œ}^*, \tag{23}$$

where M is the variance-covariance matrix of the momentum prior distribution $p(\text{Œ})$.

(iii) The updated $\mathbf{\Omega}^*$ from (23) is then used to update the the second half-step of Œ by

$$\text{Œ}^* = \text{Œ}^* + \frac{1}{2}\epsilon\frac{d}{d\mathbf{\Omega}}\log\pi(\mathbf{\Omega}|\mathbf{y})|_{\mathbf{\Omega}=\mathbf{\Omega}^*}, \tag{24}$$

3. **Accept-Reject Step:** HMC accepts the candidate $\mathbf{\Omega}^*$ as the next state of the MCMC chain, namely $\mathbf{\Omega}^{(t)}$, with probability

$$\min\left\{1, \frac{\pi(\mathbf{\Omega}^*|\mathbf{y})p(\text{Œ}^*)}{\pi(\mathbf{\Omega}^{(t-1)}|\mathbf{y})p(\text{Œ}^{(t-1)})}\right\}. \tag{25}$$

Note that we do not need to accept or reject the candidate Œ^*, because Œ is a latent variable introduced to facilitate the transition of the HMC in the parameter space. Moreover, Œ will be updated by the momentum initialization step in the next run of the HMC. The HMC finally employs the No-U-Turn strategy [50] to automatically tune parameters of ϵ and L in the leapfrog step.

3.3. Stacked Neural Networks

In this subsection, we employ the properties of artificial neural networks (NNs) to incorporate information contained in the posterior distribution of the parameters in estimating the critical collapse functions. As a state-of-the-art predictive method, NNs find the solutions to the DE system by translating the problem into optimizing a loss function in estimating the DE variables [39]. NNs fit a multi-layer perceptron to estimate the DE variables of the system. Each layer of the NNs consists of several nodes learning the curvature and form of the DE variables over the system domain. The nodes in the layers are linked by

$$N^m(\mathbf{x}, t, \text{-}) = \eta\left(\mathbf{A}^m N^{m-1}(\mathbf{x}, t, \text{-}) + \mathbf{b}^m\right), \tag{26}$$

where $N^{m-1}(\cdot)$ denotes the observed response from the m-th layer of the NNs after applying the activation function $\eta(\cdot)$ with $\mathbf{x}$ the input of the NNs, $\mathbf{A}^m$ a weight matrix of the m-th layer, $\mathbf{b}^m$ the bias of the m-th layer. Also, let - encode the set of all unknown parameters of (26). Let $\mathcal{N}(\mathbf{x}, t, \text{-})$ represent the output of the NNs in estimating the DE variables $\mathbf{g} = (g_1, \ldots, g_H)$ of (19). The solution to DE system (19) is finally derived by optimizing the loss function

$$\widehat{\mathcal{N}}(\mathbf{x}, t, \text{-}) = \arg\min_{\text{-}}\left(\mathcal{N}(\mathbf{x}, t, \text{-}) - \mathbf{g}\right)^2.$$

For more details about NNs and how they are used to estimate the DE variables, readers are referred to [39,51] and references therein.

In this paper, we first use the HMC sampler to find the posterior distribution for the parameter of the DE system. Let $\text{`}_l^*, l = 1, \ldots, L$ denote $L$ samples from the domain of the posterior distribution $\pi(\text{`}|\mathbf{y})$. Applying these samples to the NN solver, let $\widehat{\mathcal{N}}(\mathbf{x}(t|\text{`}_l^*), t, \text{-})$ denote the NN-based estimator given by $\text{`}_l^*, l = 1, \ldots, L$ from the posterior distribution. One can view the $\widehat{\mathcal{N}}(\mathbf{x}(t|\text{`}_l^*), t, \text{-})$ from a probabilistic viewpoint, as an

estimator for the DE variables $\mathbf{x}(t|\grave{}\,)$ whose probability distribution corresponds to the posterior distribution $\pi(\grave{}\,|\mathbf{y})$.

Accordingly, we propose a Bayesian model averaging strategy to stack the NNs in estimating the DE variables $\mathbf{x}(t|\grave{}\,)$. Let $\grave{}_l, l = 1, \ldots, L$ denote a set of candidate models from the training set $\mathbf{y}$. One can compute the distribution of $\widehat{\mathcal{N}}(\mathbf{x}(\mathbf{t}|\grave{}\,^*), t, \text{-})$ at some fixed space-time point t by

$$P(\widehat{\mathcal{N}}(\mathbf{x}(t|\grave{}\,), t, \text{-})|\mathbf{y}) = \sum_{l=1}^{L} P(\widehat{\mathcal{N}}(\mathbf{x}(t|\grave{}\,^l), t, \text{-})|\mathbf{y}) \pi(\grave{}\,^l|\mathbf{y}). \tag{27}$$

Thus, from (27), the posterior mean of $\widehat{\mathcal{N}}(\mathbf{x}(\mathbf{t}|\grave{}\,^*), t, \text{-})$ at some fixed space-time point t is given by

$$E(\widehat{\mathcal{N}}(\mathbf{x}(t|\grave{}\,), t, \text{-})|\mathbf{y}) = \sum_{l=1}^{L} E(\widehat{\mathcal{N}}(\mathbf{x}(t|\grave{}\,^l), t, \text{-})|\mathbf{y}) \pi(\grave{}\,^l|\mathbf{y}). \tag{28}$$

This Bayesian estimate can be viewed simply as the weighted average of the individual NN-based estimates using the weights regulated by the posterior distribution $\pi(\grave{}\,|\mathbf{y})$.

The Bayesian model averaging (28) leads to multiple proposals to stack the NN estimators. The first stacked NN-based method is to take the unweighted average of the NNs using the samples $\grave{}\,^l, l = 1, \ldots, L$ from $\pi(\grave{}\,|\mathbf{y})$ at each space-time. Although the mean stacked NNs are computed based on the unweighted average of the NNs, the random sampling of the candidates from the posterior distribution indirectly accommodates $\pi(\grave{}\,|\mathbf{y})$ in (28). Thus, the mean stacked NN estimate can be considered the posterior mean estimate. In the second approach, we propose the posterior median estimate. Here, we first stack the NNs estimates $\widehat{\mathcal{N}}(\mathbf{x}(t|\grave{}\,^l), t, \text{-}), l = 1, \ldots, L$ and then compute the median of them at each space-time point in the domain of the DE system.

4. Numerical Studies

In this section, we estimate the critical collapse functions of the equations of motion in the elliptic class of 4 dimensions in a Bayesian framework. In this framework, the parameters of the equations are treated as random variables to account for the measurement error involved in numerically estimating the parameters of the equations. We also explore the effect of this complexity on the estimation of the critical collapse functions.

Various research articles in the literature [36,39] investigated the non-linear equations of motion of the axion-dilaton system in different dimensions and for different ansatz. Hatefi et al. [37,38] used the properties of the Fourier-based and spline nonlinear smothers in predicting the critical collapse functions, respectively. Recently, Hatefi et al. [39] developed an NN-based solver to show that there is no solution in the parabolic class of black holes and [36] used a grid search root-finding method to estimate the parameters of the equations of motion numerically. According to the complex and highly nonlinear pattern of the equations, all these methods in the literature use almost a series of numerical techniques to make the equations tractable and approximate the parameters of the models. These techniques include, for instance, imposing some constraints on the equations of motions such as finiteness of $f''(z)$ as $z \to z_+$. Another constraint that one employs is the vanishing of the divergent part of $f''(z)$ that produces a complex-valued constraint at z_+. As the equations are not tractable analytically in elliptic equations of motion, for example, [33] used a numerical profile-root finding method and applied numerically a grid search discrete optimization on the coordinates of the extended parameter space of the equations of motion. To do so, the equations are simplified by approximating the target functions by the first two orders of the Taylor expansions. Applying a grid search optimization on the coordinates of the extended parameter space in the equation of motion, [33] could estimate the required parameters of the equations of motion. Removing the spurious roots of the

equations, they used parameter estimates to find the solution to the equations of motion and estimate the critical collapse functions.

Unlike [36–39], to our best knowledge, for the first time in the literature, through a probabilistic viewpoint, we develop a Bayesian method to model the measurement errors into the estimation of the critical collapse functions. Hence the framework provides a robust approach against measurement errors in parameter estimation. These Bayesian estimates incorporate information from all possible parameters' outcomes, that may occur in the numerical experiments, from the posterior distribution and accommodate this complexity in estimating critical collapse functions.

Here all the critical collapse functions are treated as DE variables of DE equations of motion 16 and 17. Since all the DE variables are numerically solved simultaneously, it is reasonable to assume that the observed DE variables have the same standard deviation σ parameter representing the variability in the numerical experiments. Therefore $\Omega = (\omega, \sigma)$ will be the set of all unknown parameters of the likelihood function under equations of motion. We first use the HMC sampler, described in Subsection 3.2, to find the posterior distribution of the likelihood parameter Ω. We used the Python package Pymc3 [52] to implement the HMC for the equations of motion. To evaluate the effect of the prior information of likelihood parameters in estimating the critical functions, we considered two different prior distributions for the parameter ω. We set the first prior distribution as Gaussian distribution with a mean 1.15 and standard deviation 0.02; that is $\pi_1(\omega) \sim N(\mu = 1.15, \sigma = 0.02)$. In contrast to the unimodal Gaussian, we assumed the second prior followed non-informative Uniform distribution between (1,1.5); that $\pi_2(\omega) \sim U(a = 1, b = 1.5)$. We also used the half-Cauchy distribution with scale parameter $\beta = 0.1$ as the prior distribution for σ to incorporate the uncertainty of the observed DE variables in the likelihood function.
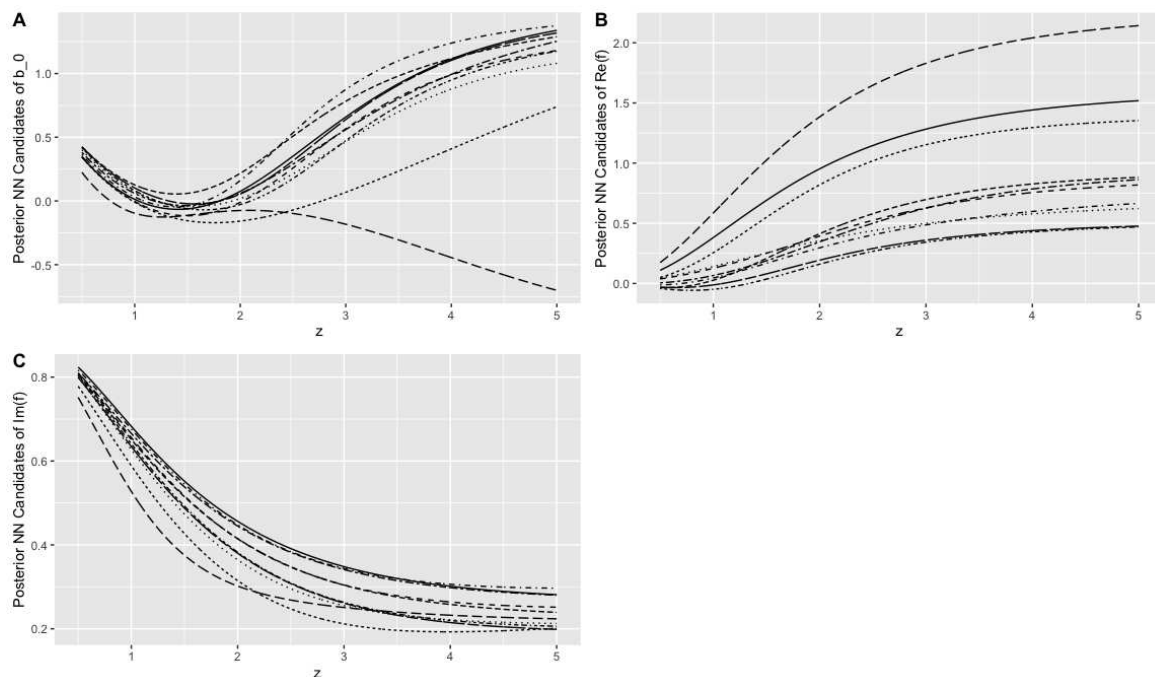
Figures 1 and 6 show the posterior distributions ω and σ under the Gaussian and Uniform prior distributions, respectively. Comparing Figures 1 and 6, under both prior distributions, we observe that the posterior distribution of ω converges to a roughly uni-modal, bell-shaped distribution where the 94% highest density interval (HDI) roughly is (1.165,1.205). In other words, considering the measurement errors from numerical analysis of the parameter estimation, the true value of ω in the equations of motion, with probability 94%, will vary between (1.165,1.205). Interestingly this is compatible with the solution for the elliptic case for four dimensions, obtained using various numerical experiments [21,34], which as reported as $\omega = 1.176$. We see that the 94% HDI contains (under both prior distributions) $w = 1.176$ as one possible outcome of the numerical analysis. It can be concluded that given the numerical measurement errors involved in the DE variables, the probability that the true value of ω appears 1.176 is roughly 13%; that is $\pi(w \in 1.176 \pm 10^{-3}|\mathbf{y}) \propto 0.13$.
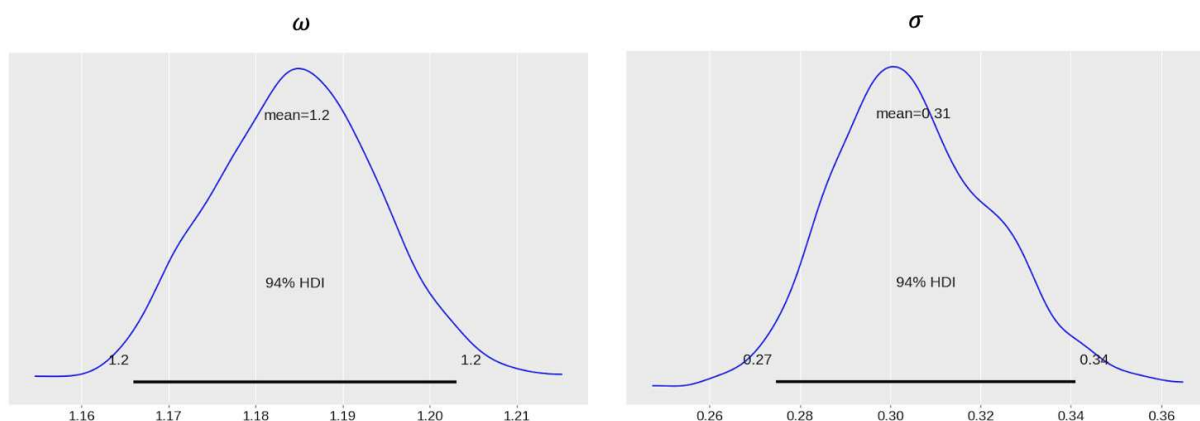
Figure 1. The posterior distribution of ω (left) and σ (right) under Gaussian prior distribution. The 94% HDIs are shown by black lines.

In the second part of the analysis, we incorporated the posterior information of the parameters into the NNs, described in Subsection 3.3, to obtain the NN-based posterior mean and posterior median estimators of the critical collapse functions $b_0(z), \text{Im}(f(z))$ and $\text{Re}(f(z))$. To do so, we randomly took $L = 200$ realizations $\omega_l^*, l = 1, \ldots, L$ from the posterior distribution $\pi(\omega|\mathbf{y})$. We then applied these 200 realizations iteratively and obtained the NN-based estimates of $b_0(z|\omega_l^*), \text{Re}(f(z|\omega_l^*))$ and $\text{Im}(f(z|\omega_l^*))$ given that $\omega = \omega_l^*, l = 1, \ldots, L$ at $n = 1000$ equally spaced $z \in [0, 5]$ in the domain of the equations of motion. These NN-based estimators are denoted by $\widehat{\mathcal{N}}\left(b_0(z_i|\omega_l^*), z, \text{-}\right), \widehat{\mathcal{N}}\left(\text{Re}(f(z_i|\omega_l^*), z, \text{-})\right)$ and $\widehat{\mathcal{N}}\left(\text{Im}(f(z_i|\omega_l^*), z, \text{-})\right)$ for $l = 1, \ldots, L$ and $i = 1, \ldots, n$. In each iteration, we implement the NNs solver using the Python Package NeuroDiffEq [53] configured with four hidden layers, each of 16 nodes, following [39]. We ran the solvers for 1000 epochs using the initial boundary conditions in Eq. 18.

Before establishing the Bayesian stacked NN estimators, we would like to explain the probabilistic perspective in sampling from the posterior distribution obtained. The posterior distribution $\pi(\omega|\mathbf{y})$ lists all possible values of ω and how often they are observed as the posterior estimate of the parameter of the equations of motion. The posterior candidate ω^*, as a sample from the posterior distribution, changes from one sample to another. Accordingly, given the posterior candidate $\omega = \omega^*$, the posterior NN candidates, as estimators for the critical collapse functions, will change from one sample to another. The probability that each of these posterior NN candidates is observed in estimating the true critical collapse functions corresponds to the posterior distribution $\pi(\omega|\mathbf{y})$. To show this probabilistic perspective, Figures 2 and 7 represent 10 posterior NN candidates of the critical collapse functions based on 10 randomly selected $\omega_l^*, l = 1, \ldots, 10$ from $L = 200$ realizations of $\pi(\omega|\mathbf{y})$ using the Gaussian and Uniform prior distributions, respectively.


Figure 2. Ten posterior NN candidates for critical collapse functions $b_0(z)$ (**A**), $\text{Re}(f(z))$ (**B**) and $\text{Im}(f(z))$ (**C**) based on ten randomly selected samples from the domain of the posterior distribution $\pi(\omega|\mathbf{y})$ under Gaussian prior distribution. Each candidate is shown by a different line type.

Because we have L Bayesian NNs candidates corresponding to L realizations from the posterior distribution, we applied the Bayesian model averaging. We obtained the two mean and median stacked NNs for all the critical functions at each space-time $z_i, i = 1, \ldots, n$. In addition to posterior mean and median estimates, we also computed the 95% Bayesian credible intervals for the critical functions on the entire

domain. To compute the Bayesian 95% credible intervals, we sorted the NN estimates from the smallest value to the largest value at each space-time. We then computed the 2.5 and 97.5 percentiles of the estimates as lower and upper bands of the interval, respectively. To compare our method with the non-Bayesian approach of [39], we computed the NNs-based estimates of the critical collapse functions similar to [39], treating $\omega = 1.176$ with the same configuration as described above. This estimate is, henceforth, called Fixed NNs.

Figures 3 and 8 show the Bayesian posterior mean and median estimates and their 95% credible intervals using Gaussian and Uniform prior distributions, respectively. Moreover, Figures 3 and 8 compare the Bayesian estimates with their corresponding fixed NN-based estimates of the critical functions. One can conclude from the figures that, given all possible values for ω in the equations of motion, the true critical collapse functions, with probability 0.95, should be in the reported Bayesian credible intervals. Also, we observe that all 95% credible intervals contain the Fixed NNs estimates of [39] as one possible candidate for estimating the critical functions. We also observe that the posterior mean and median NN estimator represent the form and curvature of the critical collapse functions as the average and median of all the Bayesian NN-based candidates.
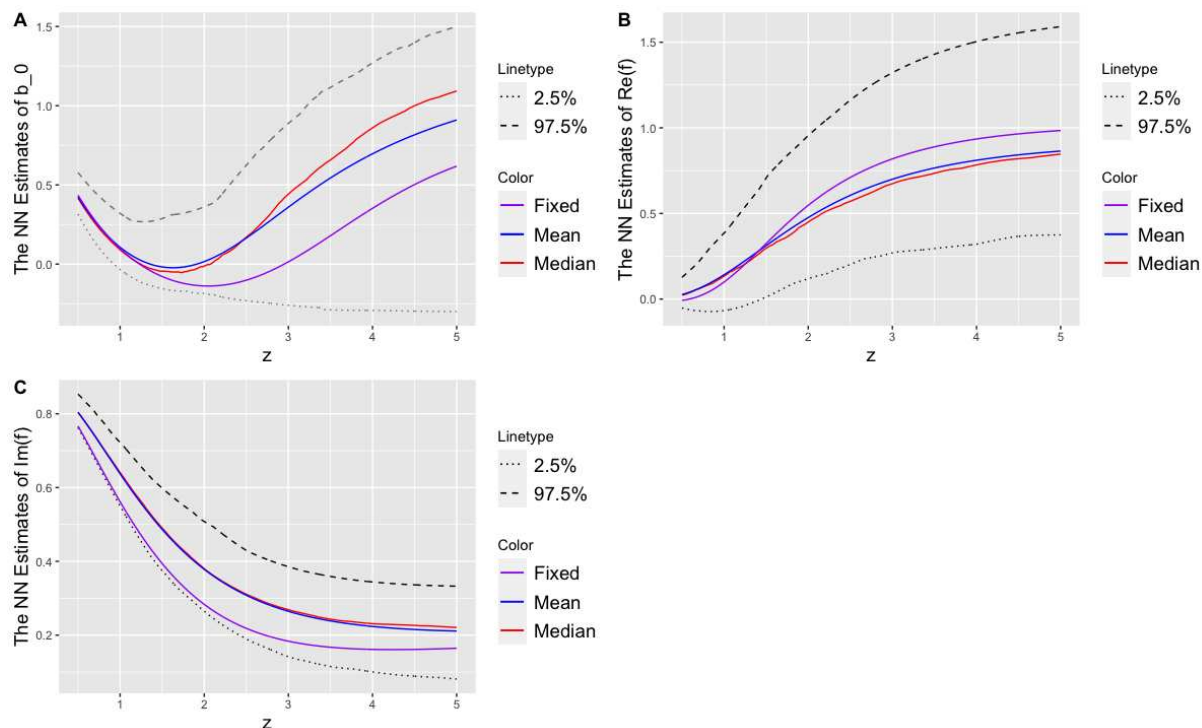

Figure 3. The mean (blue), the median (red) stacked NNs estimates along with their 95% Bayesian credible intervals for critical functions $b_0(z)$ (**A**), $\mathrm{Re}(f(z))$ (**B**) and $\mathrm{Im}(f(z))$ (**C**) under Gaussian prior distribution. The corresponding Fixed NN-based estimates are shown in purple.

Note that the fixed NN-based method of [39] is not robust against the variability and measurement error in calculating the parameter. In other words, the Fixed NN estimates will change slightly from $w = 1.176$ in estimating the parameter's value. Unlike the fixed NNs, the Bayesian credible intervals, posterior mean and median NN estimators will remain robust against changes in measurement errors in estimating w because these Bayesian estimates have already considered all possibilities of w in the domain from the posterior distribution.

Figures 4 and 9 show the posterior mean and median stacked NNs and 95% credible intervals of all the critical collapse functions together under Gaussian and Uniform prior distributions, respectively. It is

apparent from Figures 4 and 9 that the NN-based Bayesian credible intervals confirm that there must be a single self-similar solution, on average, in the area $z \in [1.5, 4.5]$ to the elliptic equations of motion in 4 dimensions where the imaginary and real parts of the G function intersects, where the explicit form of the G function for the elliptic case in 4 dimension was given in [33]. Interestingly, this finding is compatible with the finding in the literature where [36,39] showed that there was a single self-similar solution at $z = 2.56$ when the parameter of the equation is estimated to be $w = 1.176$.
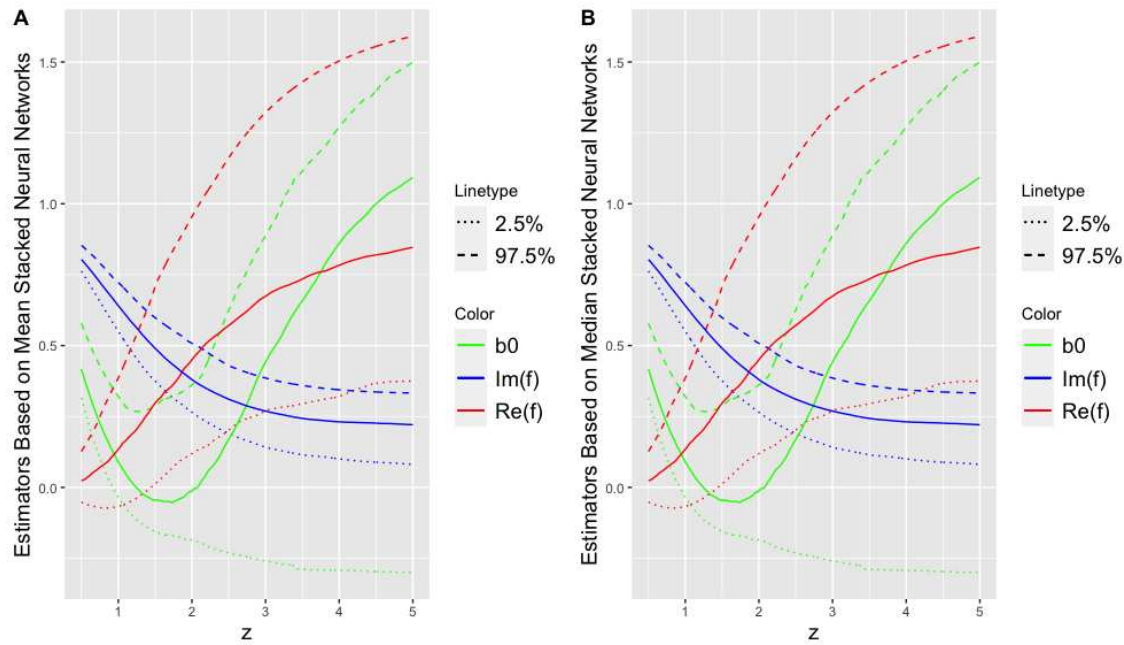
Figure 4. The mean (**A**), median (**B**) stacked NNs and their 95% credible intervals for all critical functions $b_0(z)$ (green), $\mathrm{Re}(f(z))$ (red) and $\mathrm{Im}(f(z))$ (blue) under Gaussian prior distribution.

In order to evaluate the convergence of the proposed NN-based estimates using the posterior distribution $\pi(\omega|\mathbf{y})$, we computed the train and validation losses in the last epoch of the NN solvers for each $\omega_l^*, l = 1, \ldots, L$. Figures 5 and 10 show the trace and box plots of the loss differences for $L = 200$ realizations from the posterior distribution $\pi(\omega|\mathbf{y})$ under Gaussian and Uniform prior distributions, respectively. We clearly observe that the loss differences, on average, are less than $|3 \times 10^{-3}|$ which validates the convergence of NN-based estimates in estimating the critical collapse functions in the elliptic class of 4 dimensions.
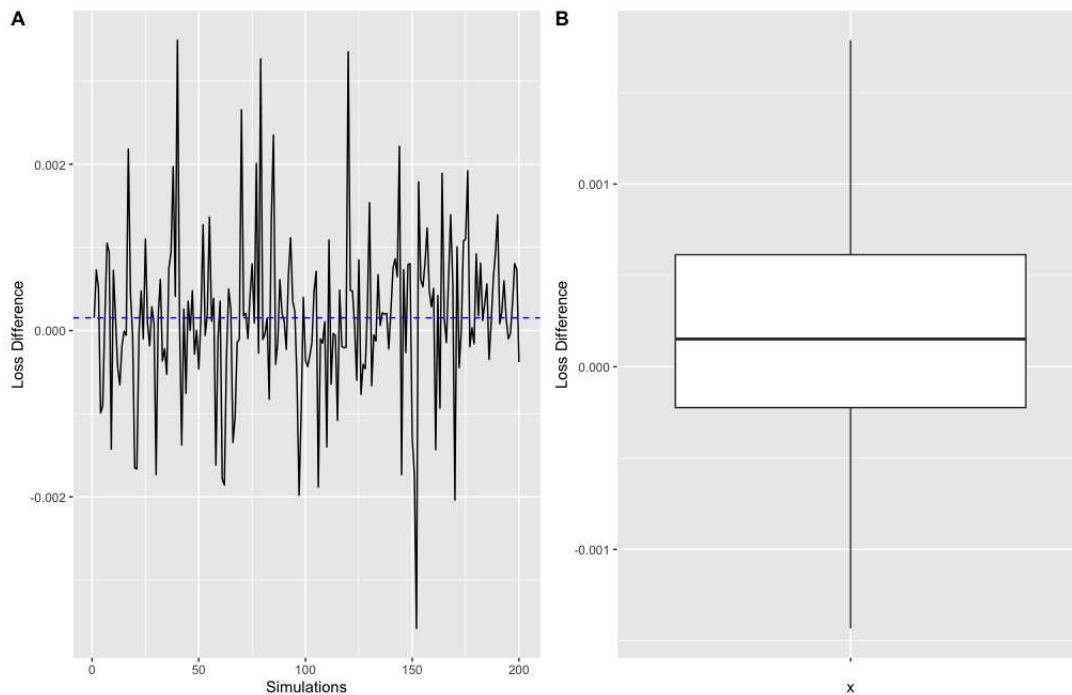
Figure 5. The trace plot (**A**) and boxplot (**B**) of the differences between train loss and validation loss in the NNs estimates using 200 realizations of the posterior distribution $\pi(\theta|\mathbf{y})$ under Gaussian prior distribution.

5. Summary and Concluding Remarks

This paper proposes a Bayesian methodology to model the self-similar solutions for the spherical gravitational collapse for an elliptic case in four dimensions. In the literature on the axion-dilaton system, due to the complex and highly nonlinear equations of motion, one needs to employ various numerical techniques and constraints to make the equations of motion tractable. These constraints include, for instance, the finiteness of $f''(z)$ as $z \to z_+$ and the vanishing of the divergent part of $f''(z)$ producing a complex-valued constraint at z_+. Because the equations of motion are not tractable analytically, for instance, [33] used a series of numerical methods to find the solutions to the equations. Their numerical profile root-finding includes, for example, approximating the main effects of the equations by the first two orders of the Taylor expansion, applying a numerical grid search discrete optimization on the coordinates of the extended parameter space of the equations and removing the spurious roots, to name a few. This enables them to estimate the parameters of the equations and then find self-similar black hole solutions. According to the complex form of the equations and the vital role of the parameters, researchers have to ignore the measurement errors imposed in finding the parameters through numerical techniques.

To fix this problem, we have proposed a novel pipeline to take the measurement errors into the statistical models in finding the solution to the elliptic equations of motion. To do so, unlike the methods in the literature, we developed a Bayesian estimation method where the parameter is treated as a random variable. We proposed Hamiltonian Monte Carlo to derive the posterior distribution of the parameter efficiently. The posterior distribution determines all the possible outcomes in estimating the parameter of the equations of motion. The posterior distribution interestingly shows that the deterministic solution available in the literature to the elliptic equations of motion, namely $\omega = 1.176$, is the true value for the parameter with a probability of almost 0.13. In addition, the posterior distribution lists all possible values for the parameter of elliptic equations of motion along with their probabilities.

Next, we utilized neural networks (NNs) to incorporate the posterior information in solving the equations of motion and estimating the critical collapse functions in the entire domain of the equations. From a probabilistic viewpoint, one can find the NN estimate of the critical function as an estimator, which

would be the true trajectory of the critical collapse function whose probability distribution corresponds to the posterior distribution of the parameter. Accordingly, we applied a Bayesian model averaging strategy and found the posterior mean and median stacked NNs and the 95% Bayesian credible intervals in estimating the critical collapse functions.

Given all variability in estimating the parameter of the equations of motion, the developed Bayesian credible intervals will contain the true critical collapse functions with a probability of 0.95. Comparing our Bayesian method with the deterministic NN approach of [39], the developed Bayesian credible intervals contain the deterministic estimate as one possible candidate in estimating the critical functions. Unlike the estimate of [39], the Bayesian proposals will remain robust against measurement errors in estimating ω because these Bayesian estimates have already considered all possibilities of the parameter in the domain of the posterior distribution. It is worth mentioning that our approach will help us to reveal the value and range of the critical exponent of axion-dilaton system in a wider range in different dimensions that we would like to pursue in the near future.

6. Appendix
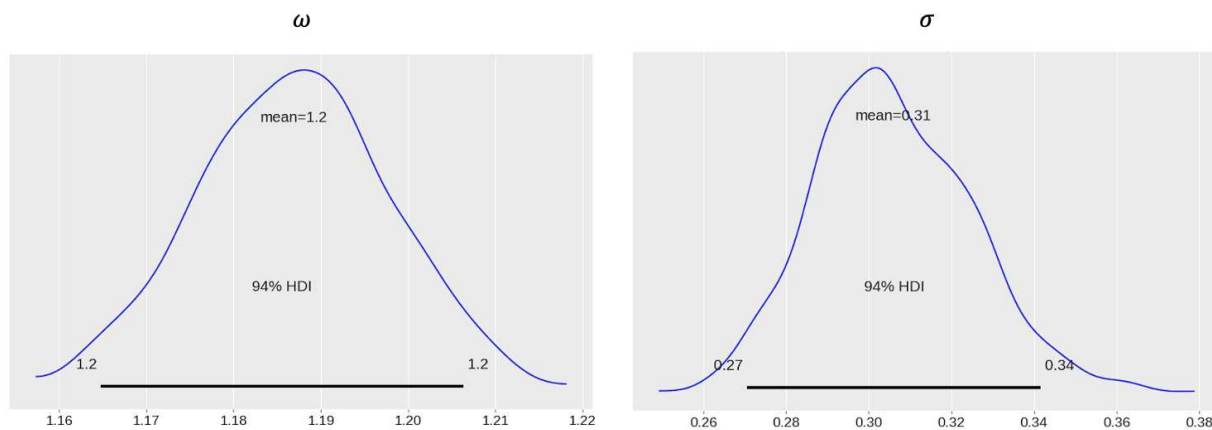
Figure 6. The posterior distribution of ω (left) and σ (right) under Uniform prior distribution. The 94% shortest Bayesian credible intervals (HDI) are shown by black lines.
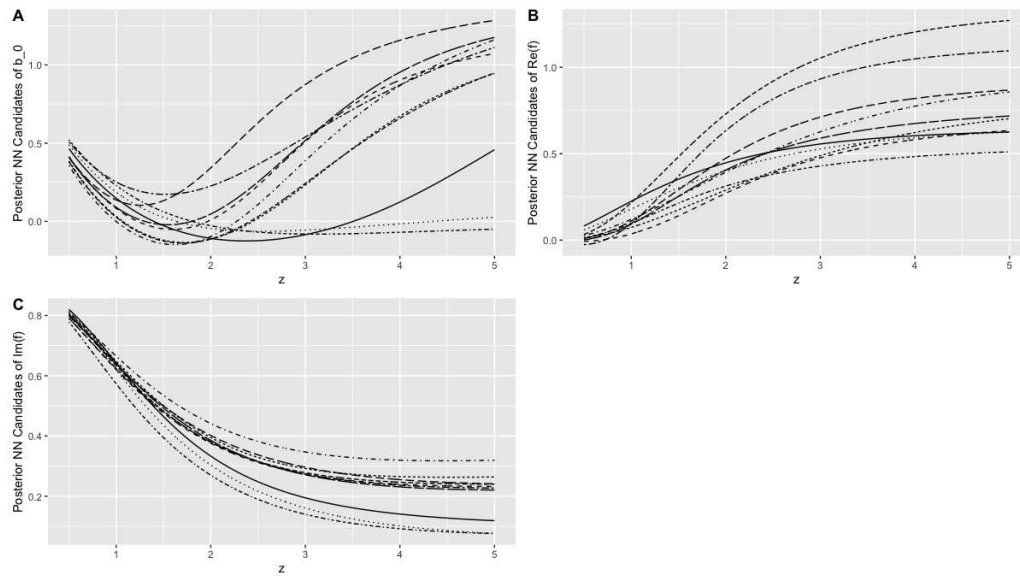
Figure 7. Ten posterior NN candidates for critical collapse functions $b_0(z)$ (**A**), $\mathrm{Re}(f(z))$ (**B**) and $\mathrm{Im}(f(z))$ (**C**) based on ten randomly selected samples from the domain of the posterior distribution $\pi(\omega|\mathbf{y})$ under Uniform prior distribution. Each candidate is shown by a different line type.
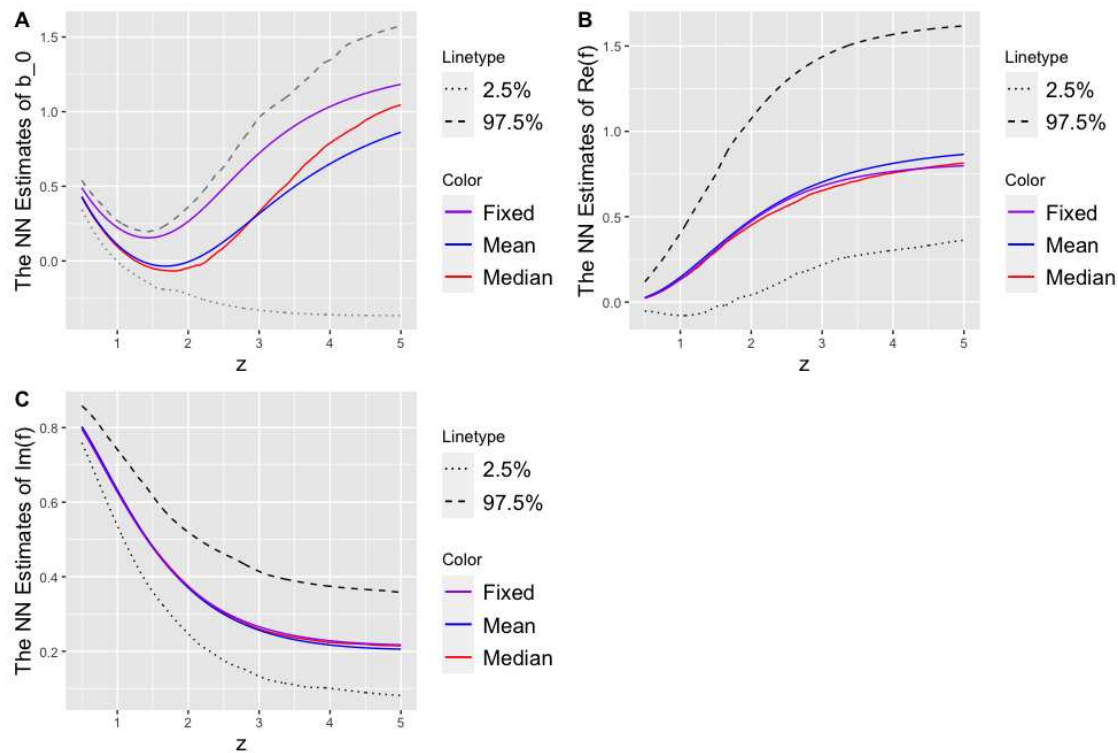
Figure 8. The mean (blue), the median (red) stacked NNs estimates along with their 95% Bayesian credible intervals for critical functions $b_0(z)$ (**A**), $\mathrm{Re}(f(z))$ (**B**) and $\mathrm{Im}(f(z))$ (**C**) under Uniform prior distribution. The corresponding Fixed ANN-based estimates are shown in purple.
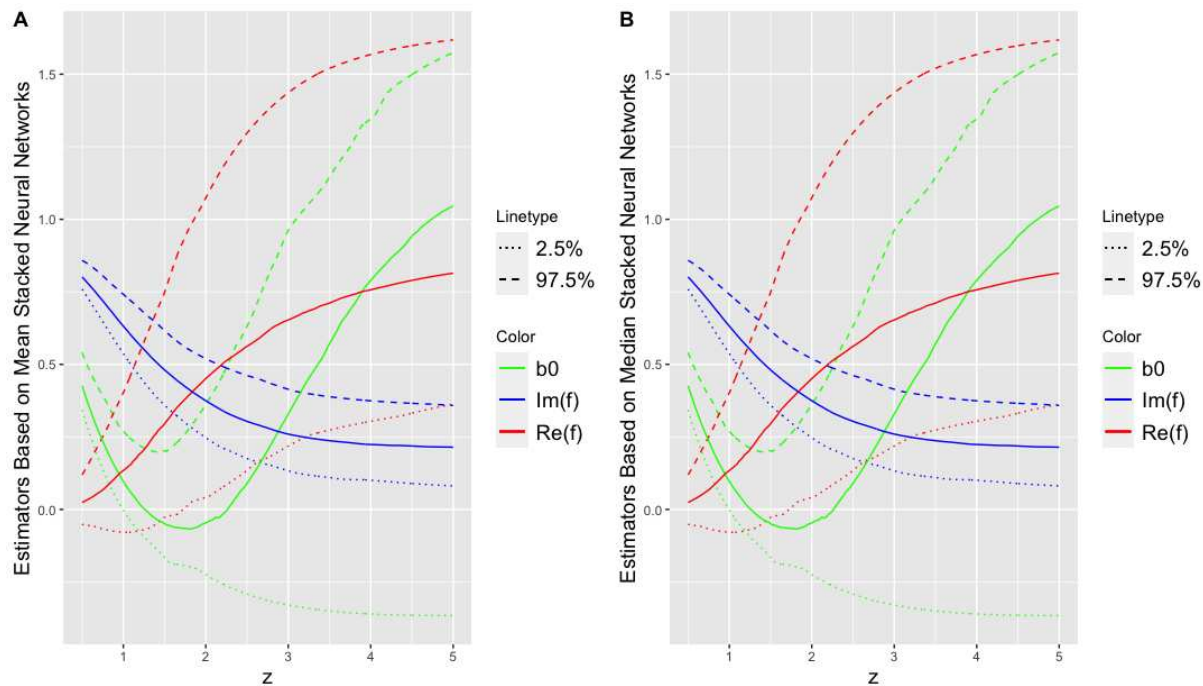
Figure 9. The mean (**A**), median (**B**) stacked NNs and their 95% credible intervals for all critical functions $b_0(z)$ (green), $\mathrm{Re}(f(z))$ (red) and $\mathrm{Im}(f(z))$ (blue) under Uniform prior distribution.
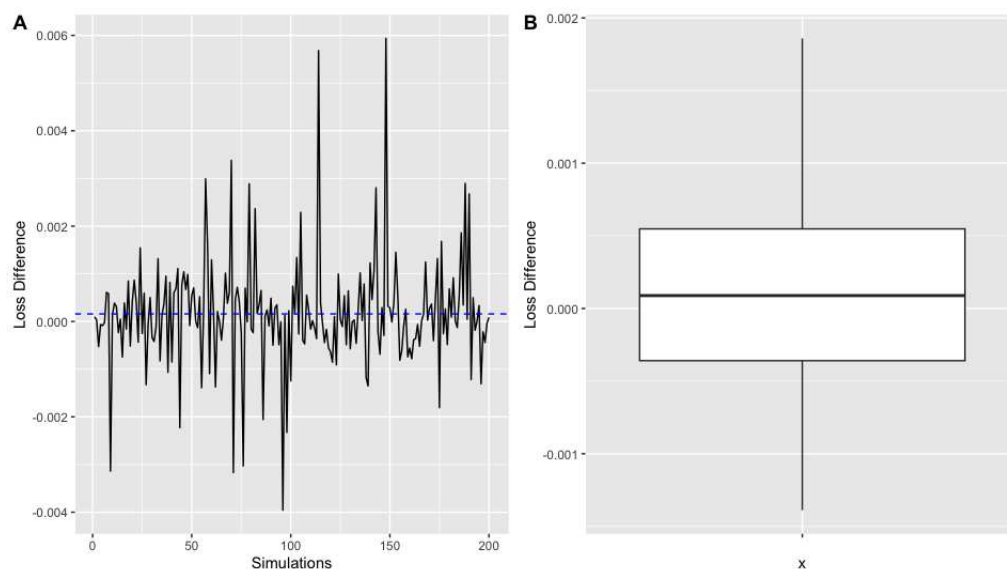
Figure 10. The trace plot (**A**) and boxplot (**B**) of the differences between train loss and validation loss in the NNs estimates using 200 realizations of the posterior distribution $\pi(\theta|\mathbf{y})$ under Uniform prior distribution.

References

1. M.W. Choptuik, Universality and Scaling in Gravitational Collapse of a Massless Scalar Field, *Phys. Rev. Lett.* **70**, 9 (1993).
2. D. Christodoulou, "The Problem of a Self-gravitating Scalar Field," Commun. Math. Phys. **105** (1986) 337; "Global Existence of Generalized Solutions of the Spherically Symmetric Einstein Scalar Equations in the Large," Commun. Math. Phys. **106** (1986) 587;"The Structure and Uniqueness of Generalized Solutions of the Spherically Symmetric Einstein Scalar Equations," Commun. Math. Phys. **109** (1987) 591.
3. R. S. Hamade, J. H. Horne and J. M. Stewart 1995 Continuous Self-Similarity and *S*-Duality, *Class. Quant. Grav.* **13** (1996) 2241-2253 [arXiv:gr-qc/9511024].

4.　R. S. Hamade and J. M. Stewart, "The Spherically symmetric collapse of a massless scalar field," Class. Quant. Grav. **13** (1996) 497 [arXiv:gr-qc/9506044].

5.　T. Koike, T. Hara, and S. Adachi, "Critical Behavior in Gravitational Collapse of Radiation Fluid: A Renormalization Group (Linear Perturbation) Analysis," Phys. Rev. Lett. **74** (1995) 5170 [gr-qc/9503007].

6.　L. Alvarez-Gaume, C. Gomez and M. A. Vazquez-Mozo, "Scaling Phenomena in Gravity from QCD," Phys. Lett. B **649** (2007) 478 [hep-th/0611312].

7.　M. Birukou, V. Husain, G. Kunstatter, E. Vaz and M. Olivier, "Scalar field collapse in any dimension," Phys. Rev. D **65** (2002) 104036 [gr-qc/0201026].

8.　V. Husain, G. Kunstatter, B. Preston and M. Birukou, "Anti-de Sitter gravitational collapse," Class. Quant. Grav. **20** (2003) L23 [gr-qc/0210011]; J. Bland, B. Preston, M. Becker, G. Kunstatter, V. Husain, Dimension dependence of the critical exponent in spherically symmetric gravitational collapse, *Class. Quant. Grav.* **22** 5355-5364

9.　E. Sorkin and Y. Oren, "On Choptuik's scaling in higher dimensions," Phys. Rev. D **71**, 124005 (2005) [arXiv:hep-th/0502034].

10.　J. Bland, B. Preston, M. Becker, G. Kunstatter and V. Husain, "Dimension-dependence of the critical exponent in spherically symmetric gravitational collapse," Class. Quant. Grav. **22** (2005) 5355 [gr-qc/0507088].

11.　E. W. Hirschmann and D. M. Eardley, "Universal scaling and echoing in gravitational collapse of a complex scalar field," Phys. Rev. D **51** (1995) 4198 [gr-qc/9412066].

12.　J. V. Rocha and M. Tomašević, "Self-similarity in Einstein-Maxwell-dilaton theories and critical collapse," Phys. Rev. D **98** (2018) no.10, 104063 [arXiv:1810.04907 [gr-qc]].

13.　L. Alvarez-Gaume, C. Gomez, A. Sabio Vera, A. Tavanfar and M. A. Vazquez-Mozo,"Critical gravitational collapse: towards a holographic understanding of the Regge region," Nucl. Phys. B **806** (2009) 327 [arXiv:0804.1464 [hep-th]].

14.　C. R. Evans and J. S. Coleman, "Observation of critical phenomena and selfsimilarity in the gravitational collapse of radiation fluid," Phys. Rev. Lett. **72** (1994) 1782 [gr-qc/9402041].

15.　D. Maison, "Non-Universality of Critical Behaviour in Spherically Symmetric Gravitational Collapse," Phys. Lett. B **366** (1996) 82 [gr-qc/9504008].

16.　A. Strominger and L. Thorlacius, "Universality and scaling at the onset of quantum black hole formation," Phys. Rev. Lett. **72** (1994) 1584 [hep-th/9312017].

17.　E. W. Hirschmann and D. M. Eardley, 'Critical exponents and stability at the black hole threshold for a complex scalar field," Phys. Rev. D **52** (1995) 5850 [gr-qc/9506078].

18.　A. M. Abrahams and C. R. Evans, "Critical behavior and scaling in vacuum axisymmetric gravitational collapse," Phys. Rev. Lett. **70** (1993) 2980.

19.　L. Alvarez-Gaume, C. Gomez, A. Sabio Vera, A. Tavanfar and M. A. Vazquez-Mozo, "Critical formation of trapped surfaces in the collision of gravitational shock waves," JHEP **0902**, 009 (2009) [arXiv:0811.3969 [hep-th]].

20.　E. W. Hirschmann and D. M. Eardley, "Criticality and bifurcation in the gravitational collapse of a selfcoupled scalar field," Phys. Rev. D **56** (1997) 4696 [gr-qc/9511052].

21.　D. M. Eardley, E. W. Hirschmann and J. H. Horne,1995 S duality at the black hole threshold in gravitational collapse, *Phys. Rev.* D **52** (1995) 5397 (arXiv:gr-qc/9505041)

22.　E. Hatefi and E. Vanzan, "On higher dimensional self-similar axion–dilaton solutions," Eur. Phys. J. C **80** (2020) no.10, 952 [arXiv:2005.11646 [hep-th]].

23.　E. Hatefi and A. Kuntz, "On perturbation theory and critical exponents for self-similar systems," Eur. Phys. J. C **81** (2021) no.1, 15 [arXiv:2010.11603 [hep-th]].

24.　Maldacena J M 1998 The large N limit of superconformal field theories and supergravity *Adv. Theor. Math. Phys.* **2** 23152 (arXiv:hep-th/9711200)

25.　Witten E 1998 Anti-de Sitter space and holography *Adv. Theor. Math. Phys.* **2** 25391 (arXiv:hep-th/9802150)

26.　Gubser S S, Klebanov I R and Polyakov A M 1998 Gauge theory correlators from non-critical string theory *Phys. Lett.* **B 428** 10514 (arXiv:hep-th/9802109)

27.　Witten E 1998 Anti-de Sitter space, thermal phase transition, and confinement in gauge theories *Adv. Theor. Math. Phys.* **2** 50532 (arXiv:hep-th/9803131)

28.　D. Birmingham, "Choptuik scaling and quasinormal modes in the AdS / CFT correspondence," Phys. Rev. D **64** (2001), 064024 [arXiv:hep-th/0101194 [hep-th]].

29.　L. Álvarez-Gaumé, C. Gómez, and M. A. Vázquez-Mozo, 2007 Scaling Phenomena in Gravity from QCD *Phys. Lett.* **B649** 478-482 (arXiv: hep-th/0611312)

30. E. Hatefi, A. Nurmagambetov and I. Park, "ADM reduction of IIB on $\mathcal{H}^{p,q}$ to dS braneworld," JHEP **04** (2013), 170, arXiv:1210.3825 , "N^3 entropy of M5 branes from dielectric effect," Nucl. Phys. B **866** (2013), 58-71, arXiv:1204.2711, S. de Alwis, R. Gupta, E. Hatefi and F. Quevedo, "Stability, Tunneling and Flux Changing de Sitter Transitions in the Large Volume String Scenario," JHEP **11** (2013), 179, arXiv:1308.1222.

31. A. Ghodsi and E. Hatefi, "Extremal rotating solutions in Horava Gravity," Phys. Rev. D **81** (2010) 044016 [arXiv:0906.1237 [hep-th]].

32. R. S. Hamade, J. H. Horne and J. M. Stewart, "Continuous Self-Similarity and S-Duality," Class. Quant. Grav. **13** (1996) 2241 [arXiv:gr-qc/9511024].

33. R. Antonelli and E. Hatefi, "On self-similar axion-dilaton configurations," , JHEP **03** (2020), 074 [arXiv:1912.00078 [hep-th]].

34. L. Álvarez-Gaumé and E. Hatefi, "Critical Collapse in the Axion-Dilaton System in Diverse Dimensions," Class. Quant. Grav. **29** (2012) 025006 [arXiv:1108.0078 [gr-qc]].

35. L. Álvarez-Gaumé and E. Hatefi, "More On Critical Collapse of Axion-Dilaton System in Dimension Four," JCAP **1310** (2013) 037 [arXiv:1307.1378 [gr-qc]].

36. R. Antonelli and E. Hatefi, "On Critical Exponents for Self-Similar Collapse," JHEP **03** (2020), 180 [arXiv:1912.06103 [hep-th]].

37. E. Hatefi and A. Hatefi, "Estimation of Critical Collapse Solutions to Black Holes with Nonlinear Statistical Models," Mathematics **10** (2022) no.23, 4537 [arXiv:2110.07153 [gr-qc]].

38. E. Hatefi and A. Hatefi, "Nonlinear statistical spline smoothers for critical spherical black hole solutions in 4-dimension," Annals Phys. **446** (2022), 169112 [arXiv:2201.00949 [gr-qc]].

39. E. Hatefi, A. Hatefi and R. J. López-Sastre,"Analysis of Black Hole Solutions in Parabolic Class Using Neural Networks," Eur. Phys. J. C (2023) 83:623 [arXiv:2302.04619 [gr-qc]].

40. Original reviews are A. Sen,"Strong - weak coupling duality in four-dimensional string theory," Int. J. Mod. Phys. A 9, 3707 (1994) [hep-th/9402002]; J.H. Schwarz, "Evidence for non perturbative string symmetries," Lett. Math. Phys. 34, 309 (1995)[hep-th/9411178].

41. M.B. Green, J.H. Schwarz and E. Witten, 1987 *Superstring Theory* Vols I,II, Cambridge University Press,

42. J. Polchinski, 1998 *String Theory*, Vols I,II, Cambridge University Press

43. A. Font, L. E. Ibanez, D. Lust and F. Quevedo, "Strong - weak coupling duality and nonperturbative effects in string theory," Phys. Lett. B **249** (1990) 35.

44. Girolami, M. (2008). Bayesian inference for differential equations. Theoretical Computer Science, 408(1), 4-16.

45. Robert, C. P., Casella, G. (1999). Monte Carlo statistical methods (Vol. 2). New York: Springer.

46. Robert, C. P. (2007). The Bayesian choice: from decision-theoretic foundations to computational implementation (Vol. 2). New York: Springer.

47. Bishop, C. M. (2006). Pattern recognition and machine learning. New York: Springer.

48. Neal, R. M. (2011). MCMC using Hamiltonian dynamics. Handbook of markov chain monte carlo, 2(11), 2.

49. Betancourt, M., & Girolami, M. (2015). Hamiltonian Monte Carlo for hierarchical models. Current trends in Bayesian methodology with applications, 79(30), 2-4.

50. Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15(1), 1593-1623.

51. Goodfellow, I., Bengio, Y., & Courville, A. "Deep learning", MIT press. (2016)

52. Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. PeerJ Computer Science, 2, e55.

53. F. Chen, D. Sondak, P. Protopapas, P.Mattheakis, M. Liu, S.Agarwal, D. & Di Giovanni, M.,"NeuroDiffEq: A Python package for solving differential equations with neural networks," Journal Of Open Source Software. **5**, 1931 (2020)