THE EUROPEAN
PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Sequential Monte Carlo with cross-validated neural networks for complexity of hyperbolic black hole solutions in 4D

Armin Hatefi[1,a], Ehsan Hatefi[2,b]

[1] Department of Mathematics and Statistics, Memorial University of Newfoundland, St John's, NL, Canada
[2] Research Group GRAM, Department of Signal Theory and Communications, University of Alcalá, Alcalá de Henares, Spain

Abstract This paper investigates the self-similar solutions of the Einstein-axion-dilaton configuration from type IIB string theory and the global SL(2,R) symmetry. We consider the Continuous Self Similarity (CSS), where the scale transformation is controlled by an SL(2, R) boost or hyperbolic translation. The solutions stay invariant under the combination of space-time dilation with internal SL(2,R) transformations. We develop a new formalism based on Sequential Monte Carlo (SMC) and artificial neural networks (NNs) to estimate the self-similar solutions to the equations of motion in the hyperbolic class in four dimensions. Due to the complex and highly nonlinear patterns, researchers typically have to use various constraints and numerical approximation methods to estimate the equations of motion; thus, they have to overlook the measurement errors in parameter estimation. Through a Bayesian framework, we incorporate measurement errors into our models to find the solutions to the hyperbolic equations of motion. It is well known that the hyperbolic class suffers from multiple solutions where the critical collapse functions have overlap domains for these solutions. To deal with this complexity, for the first time in literature on the axion-dilaton system, we propose the SMC approach to obtain the multi-modal posterior distributions. Through a probabilistic perspective, we confirm the deterministic α and β solutions available in the literature and determine all possible solutions that may occur due to measurement errors. We finally proposed the penalized Leave-One-Out Cross-validation (LOOCV) to combine the Bayesian NN-based estimates optimally. The approach enables us to determine the optimum weights while dealing with the co-linearity issue in the NN-based estimates and better predict

the critical functions corresponding to multiple solutions of the equations of motion.

1 Introduction

It is well known that all Black holes can be characterised by their mass, their charge as well as their angular momentum. Choptuik in [1] also showed that there is yet one more parameter that explains the critical gravitational collapse solutions and it is called the critical exponent. More specifically, Christodolou in [2–4] first had revealed the spherically symmetric collapse of the real scalar field. Later on, Choptuik [1] numerically showed that the real scalar field gravitational collapse solution demonstrates the discrete self-similarity property. Indeed, the gravitational solution shows space-time self-similarity where the dilations can take place. Therefore, the critical solution does provide an scaling law. If we show the initial condition of the real scalar field by parameter p, which is called the field amplitude, then $p = p_{\text{crit}}$ defines the critical solution and hence, the black hole can be formed once p chooses the values bigger than p_{crit}. Indeed, for $p > p_{\text{crit}}$ the mass of the black hole or the Schwarzschild radius are given by a scaling law as follows

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

The following articles [1,5,6] found that the critical exponent for a real scalar field is given by $\gamma \simeq 0.37$ in four dimensions. Notice that for dimensions bigger than four ($d \geq 4$), the black hole's mass scaled by [7,8] as

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

One may read some other numerical investigations for several other matter content in [9–14]. The collapse solutions of the perfect fluid had been studied in [7,15–17] and its critical exponent $\gamma \simeq 0.36$ was also found in [16]. The authors in

[a] e-mail: ahatefi@mun.ca

[b] e-mails: ehsan.hatefi@uah.es; ehsanhatefi@gmail.com (corresponding author)

[18] argued that γ might have a universal value for all matter fields that can be coupled to gravity in four dimensions. As discussed in [7,17,19] the critical exponent is explored by using the perturbations of self-similar solutions.

The solutions with axial symmetry were investigated in [20], while shock waves are studied in [21]. The authors in [22] explored the value of the critical exponent $\gamma \simeq 0.2641$ for the axion-dilaton configuration in four dimensions. Interestingly, in [23] we have studied the perturbations and were able to precisely generate the existing value [22] of $\gamma \sim 0.2641$ in four dimensions and other critical exponents have been derived in [24] in four and five dimensions.

Albeit the authors in [5,25] investigated the entire analysis for the elliptic case in four dimensions, it is worth mentioning that their methods can also be examined in hyperbolic and parabolic cases as well as other dimensions, where for further results we refer to [24,26].

Let us provide various motivation for the study of the critical collapse of the Einstein-axion-dilaton configuration. The first motivation is related to the gauge-gravity duality [27–30], corresponding the choptuik exponent, the imaginary part of quasi-normal modes as well as the dual conformal field theory that is pointed out in [31]. In fact, of the interest is to study the spaces that approach asymptotically to $AdS_5 \times S^5$, where the simplest system is to consider in type IIB string theory of the axion-dilaton system with the self-dual 5-form field. Then, one can start to analyze the black hole solutions in diverse dimensions. It is also important to highlight that this system has also been related to the holographic description of black hole formation, see [15,32]. Finally the implications of this system to black hole physics have been carried out in [33–36]. The key role of S-duality for these self similar solutions has also been considered in [37]. In this paper, we are dealing with the collapse of matter to form small mass black holes and hence one considers a small space-time region just close to where the singularity occurs. It has also been shown that this event is independent of the asymptotic structure of the space-time to which the collapse happens. In fact there is already the numerical evidence in asymptotically AdS space-times confirming that this is the case [10,38]. Therefore we eliminate the cosmological constant and just analyse self-similar solutions for the axion-dilaton system.

The self-similar solutions for all elliptic, hyperbolic and parabolic classes of SL(2,R) have been discovered in [39] in four and five dimensions for all classes of SL(2,R), which are the extensions of the earlier results [40,41]. In [42] we have also recently made use of the Fourier-based regression models for obtaining the critical solutions. Consequently, the challenges of [42] have been addressed in [43] where we applied truncated power basis, natural spline and penalized B-spline regression models accordingly in order to be able to explore the non-linear functions. In [44] we applied artificial neural networks in order to address the instability of

the black hole solutions for the specific parabolic class in higher dimensions. Lastly, in [45], we actually proposed a new formalism to be able to model the complexity of elliptic black hole solution in four dimension using hamiltonian monte carlo with stacked neural networks.

In this paper, we propose a new formalism based on Sequential Monte Carlo (SMC) and artificial neural networks (NNs) to be able to model the hyperbolic class of the spherical gravitational self-similar solutions in four dimensions. Due to the nature of highly non linear equations of motions of hyperbolic black holes, various authors used a variety of numerical calculations to simplify the equations of motions and parameters of the theory, for instance one can see [26,43,44]. Hence, due to this reason the authors must have overlooked the measurement errors that are imposed in exploring the parameters through various numerical methods.

Thus here we propose a new method to carry out the measurement errors, involved in parameter estimation, into our statistical models in exploring the solutions to the equations of motion. Recently Hatefi et al. [45] applied the Hamiltonian Monte Carlo method to find solutions to the equations of motion in the elliptic class of four dimensions in a Bayesian framework. Unlike [45], the hyperbolic equations of motion in four dimensions have multiple solutions and therefore the collapse functions do have overlap domains under these solutions. In order to deal with this challenge, for the first time in the literature on the axion-dilaton system, we proposed the SMC approach to derive the posterior distribution of the parameters. The posterior distribution does provide all possible solutions in estimating the parameter of the equations of motion. Interestingly, the posterior distribution confirms the deterministic α and β solutions found in the literature for the hyperbolic class in four dimensions. Unlike other methods in the literature, in this paper, we impose the l_2 penalized Leave-One-Out Cross-validation (LOOCV) to optimally combine the Bayesian NNs candidates. The advantage of this approach is that it also enables us to determine the optimum weights while dealing with the co-linearity issue in the NN-based estimates and better predict the critical functions corresponding to multiple solutions of the equations of motion in the hyperbolic class.

The organization of the paper is as follows. In Sect. 2 we briefly explain the relevant effective action for the axion-dilaton configuration, its equations of motion as well as the initial conditions that come from the continuous self-similarity requirement. We the describe our methodology to actually model the complexity of Sequential Monte Carlo with Cross-validated Neural Networks for hyperbolic black hole solutions in four dimensions. In Sect. 4, we use SMC samples from the posterior distribution and construct NN estimates based on the posterior mean and LOOCV as well as the 95% credible intervals in estimating the critical collapse functions corresponding to multiple solutions of the

equations of motion. Lastly, we present the results and conclude in the Sect. 5.

2 The Einstein-axion-dilaton system and its equations of motion for hyperbolic class

The two real scalar fields of axion and dilaton can be combined to construct a single complex scalar field $\tau \equiv a + ie^{-\phi}$. Its dynamics and coupling to the gravity or its effective action for four-dimensional axion-dilaton (a, Φ) system is described by

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left(R - \frac{1}{2} \frac{\partial_a \tau \partial^a \bar{\tau}}{(\mathrm{Im}\,\tau)^2} \right), \tag{3}$$

where R is the scalar curvature. If we take variations from the metric and τ then one would be able to find out all the equations of motion as follows

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

$$\nabla^a \nabla_a \tau + \frac{i \nabla^a \tau \nabla_a \tau}{\mathrm{Im}\tau} = 0. \tag{5}$$

This effective action is classically invariant under SL(2,R) transformations which means that if τ gets replaced by

$$\tau \rightarrow \frac{a\tau + b}{c\tau + d}, \tag{6}$$

where $(a, b, c, d) \in R$, $ad - bc = 1$ then g_{ab} and the action remains invariant. As argued originally by [46–50]) this group gets broken to SL(2,Z) as an indication of duality transformation.

The spherically symmetric metric is represented by [25]

$$ds^2 = (1 + u(t, r)) \left(-b(t, r)^2 dt^2 + dr^2 \right) + r^2 d\Omega_{d-2}^2. \tag{7}$$

If we take into account the time scaling for (7), (as shown in [25]), one can set $b(t, 0) = 1$ and regularity condition indicates $u(t, 0) = 0$. The so called continuous self-similarity (CSS) means there exist a killing vector ξ that generates global scale transformation, where in spherical coordinates, we define $\xi = t\,\partial/\partial t + r\,\partial/\partial r$. The assumption of continuous scale invariance for the metric gets related a scaling for the line element under dilations as follows

$$(t, r) \rightarrow (\Lambda t, \Lambda r), \quad \Lambda > 0 \tag{8}$$

then

$$ds^2 \rightarrow \Lambda^2 ds^2. \tag{9}$$

Now if we consider the scale invariant variable $z = -r/t$ then the self-similarity of the metric implies that the functions $u(t, r)$, $b(t, r)$ must be expressed in terms of z, that is

$$u(t, r) = u(z), \quad b(t, r) = b(z), \tag{10}$$

The scalar τ must also be invariant up to an SL(2, $\mathbb{R}$) transformation, so that

$$\tau(t, r) \rightarrow M(\Lambda)\tau(t, r). \tag{11}$$

Hence a system of (g, τ) that satisfies Eqs. (10), (11) to be continuously self-similar (CSS). Hence, physically distinct cases are related to the different conjugacy classes of $\frac{dM}{d\Lambda}\big|_{\Lambda=1}$.

The effective action in (3) is SL(2,R)-invariant, hence we can consider a compensation of the scale transformation of (t, r) by an SL(2,R) transformation. In fact in [40] we already found out three different possible assumptions for this particular system. Those were called the elliptic, hyperbolic and parabolic classes that are related to three classes of $SL(2, R)$ transformations which were used to compensate for a scaling transformation in space-time. Let us describe the hyperbolic ansätze for $\tau(t, z)$.

The general form of the ansatz for the hyperbolic class is given by

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

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

$$f(z) \rightarrow e^\lambda f(z), \quad \lambda \in \mathbb{R}. \tag{13}$$

Note that under $SL(2, R)$-transformation the following

$$\tau(t, r) \rightarrow (-t)^\omega f(z) \tag{14}$$

exactly produces the same equations of motion for hyperbolic case, where $f(z)$ is a complex function satisfying $\mathrm{Im}\, f(z) > 0$, and ω is a real constant. Let us describe the derivation of the equations of motion for the hyperbolic class in four dimensions. If we apply continuous self-similarity ansätze (12) to all the equations of motion (4) and (5) then one would be able to explore the ordinary differential equations for $u(z)$, $b(z)$, $f(z)$. However, if we make use of the spherical symmetry then one reveals that $u(z)$ and its first derivation $u'(z)$ can be expressed in all the equations in terms of $b(z)$, $f(z)$ and their first derivatives as follows

$$u(z) = \frac{z b'(z)}{b(z)} \tag{15}$$

$$\frac{u'(z)}{(1 + u(z))} = \frac{w \bar{f}(z) f'(z) + w f(z) \bar{f}'(z) - 2z \bar{f}'(z) f'(z)}{(f(z) - \bar{f}(z))^2} \tag{16}$$

All the ordinary differential equations (ODEs) are given by

$$b'(z) = B(b(z), f(z), f'(z)), \tag{17}$$

$$f''(z) = F(b(z), f(z), f'(z)). \tag{18}$$

Finally, the equations of motion in the hyperbolic class in four dimensions are represented by

$$b' = \frac{z(b^2 - z^2)}{b(f - \bar{f})^2} f' \bar{f}' - \frac{\omega(b^2 - z^2)}{b(f - \bar{f})^2} (f\bar{f}' + \bar{f}f') - \frac{\omega^2 z |f|^2}{b(f - \bar{f})^2} \tag{19}$$

$$
\begin{aligned}
f'' &= -\frac{z(b^2 + z^2)}{b^2(f - \bar{f})^2} f'^2 \bar{f}' + \frac{2}{(f - \bar{f})} \left(\frac{1}{\bar{f}} + \frac{\omega(b^2 + z^2)}{2b^2(f - \bar{f})} \right) \bar{f} f'^2, \\
&\quad + \frac{\omega(b^2 + 2z^2)}{b^2(f - \bar{f})^2} f f' \bar{f}' + \frac{2}{z} \left(-1 + \frac{\omega z^2 (f + \bar{f})}{(b^2 - z^2)(f - \bar{f})} \right. \\
&\quad + \left. \frac{\omega^2 z^4 |f|^2}{b^2(b^2 - z^2)(f - \bar{f})^2} \right) f' - \frac{\omega^2 z}{b^2(f - \bar{f})^2} f^2 \bar{f}' \\
&\quad + \frac{2\omega}{(b^2 - z^2)} \left(-\frac{1}{2} - \frac{\omega(f + \bar{f})}{2(f - \bar{f})} - \frac{\omega^2 z^2 |f|^2}{2b^2(f - \bar{f})^2} \right) f. \tag{20}
\end{aligned}
$$

The equation of motion for b is a first-order linear inhomogeneous equation with initial condition $b(0) = 1$. The initial conditions for $f(z)$, $f'(z)$ are also realised by applying the smoothness of the critical solution. From (20), we encounter five singularities at $z = \pm 0$, $z = \infty$ and $z = z_\pm$. One can readily show that the singularities $z = \pm 0$, $z = \infty$ can be eliminated by the coordinate transformation as shown in [41]. The last two singularities are demonstrated by $b(z_\pm) = \pm z_\pm$. They correspond with the horizon where $z = z_+$ is just a coordinate singularity as shown in [22,40]. Therefore by definition, τ must also be regular across it.

Hence $f''(z)$ must also be finite as $z \to z_+$. Therefore, one finds that the vanishing of the divergent part of $f''(z)$ produces a complex-valued constraint at z_+, that is indicated by $G(b(z_+), f(z_+), f'(z_+)) = 0$ where the explicit form of the G function for the hyperbolic case in four dimension is given in

$$
\begin{aligned}
&G(f(z_+), f'(z_+)) \\
&= \bar{f}(z_+) \left(2z_+ \left(2\omega^2 \right) f'(z_+) + 2(\omega - 1)\omega \bar{f}(z_+) \right) \\
&\quad + f(z_+) \left(2z_+(-2 + 2\omega + 2) f'(z_+) + 2\omega \bar{f}(z_+) \left(2 - \omega^2 \right) \right) \\
&\quad - \frac{2z_+ \bar{f}(z_+)^2 (2 + 2\omega - 2) f'(z_+)}{f(z_+)} - 2\omega(\omega + 1) f(z_+)^2. \tag{21}
\end{aligned}
$$

One finds it convenient to make use of change of variables as $f(z) = u(z) + iv(z)$. Indeed, using regularity at $z = 0$ and some residual symmetries one gets

$$b(0) = 1, \qquad f(0) = 1 + ix_0 \qquad (x_0 > 0) \tag{22}$$

as well as

$$f'(0) = u'(0) = v'(0) = 0 \tag{23}$$

These equations are invariant under a constant scaling $f \to \lambda f$, therefore one has the freedom to choose either the real value of $f(z)$ or its imaginary part as one wishes at a particular value of z, so we would like to set $u(0) = 1$. If one require the regularity at the origin $z = 0$, then one explores the following initial conditions for the hyperbolic class as

$$b(0) = u(0) = 1, \qquad u'(0) = v'(0) = 0 \tag{24}$$

Therefore, the real and imaginary parts of G must vanish which determines ω, where $(x_0 > 0)$ and x_0 is a real parameter. The three discrete solutions in four dimensions were explored in [23] where these solutions are found by integrating numerically the equations of motion. The solutions in hyperbolic class are identified by seeing the very rapidly decreasing $\mathrm{Im} f(0)$. No solutions are explored with $\mathrm{Im} f(0) > 1$. The α and β solutions are evidently explored. The α solution is given by

$$\omega = 1.362, \qquad v(0) = 0.708, \qquad z_+ = 1.440$$

The β solution is

$$\omega = 1.003, \qquad v(0) = 0.0822, \qquad z_+ = 3.29$$

However for the γ solution we notice that due to the fact that $\mathrm{Im} f(0)$ is so small also the z_+ root-finding gets effected with numerical noise ($\omega = 0.541$, $v(0) = 0.0059$, $z_+ = 8.44$), and hence the quality is not perfect because the $G \sim 10^{-7}$ comparing to $G \sim 10^{-13}$–10^{-17} for the first two solutions, hence we concentrate on α, β solutions.

It is also important to highlight the fact that in [24] we have shown that the Choptuik exponent depends on the dimensions, matter content as well as the different branches of the unperturbed self similar solutions. Thus, we argue that the conjecture about the universality of Choptuik exponent is not satisfied. However, we also claim that there may exist some other universal behaviours that could have been hidden in combinations of critical exponents or there might be some other parameters of the given theory that have not been considered yet in our understanding of current investigations.

3 Statistical methods

In this section, we investigate Bayesian solutions to the parameters of the hyperbolic equations of motion using sequential Monte Carlo methodology. Suppose $\mathbf{x}(t) = (x_1(t), \ldots, x_H(t))$ represent the solutions to

$$\frac{d}{dt} x_i(t) = g_i(\mathbf{x}(t)|\boldsymbol{\theta}), \tag{25}$$

where the system encompasses H differential equations (DEs) in which t denotes the space-time, θ encodes the set of all unknown parameters of the system and $x_i(t)$ denotes

the solution to i-th DE, $i = 1, \ldots, H$. Henceforth, we call $\mathbf{x}(t)$ the DE variables.

As the equations of motion in black holes are highly non-linear, typically there is no closed-form solution for the true trajectory of the DE variables. Therefore, researchers often apply a series of numerical methods and observe the trajectory of the DE subject to measurement errors. Let y_{ij} denote the observed trajectory of the i-th DE variable at $j = 1, \ldots, n_i$ space-time points. To take into account the uncertainty involved in the observed DE variables, let y_{ij} follow a Gaussian distribution with mean $x_i(t_j|\boldsymbol{\theta})$ and standard deviation σ_i for $i = 1, \ldots, H$. Let $\boldsymbol{\Omega} = (\boldsymbol{\theta}, \boldsymbol{\sigma})$ denote the set of all unknown parameters of the model where $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_H)$. Combing all information from the observed DEs, the likelihood function of $\boldsymbol{\Omega}$ is given by

$$p(\boldsymbol{\Omega}|\mathbf{y}) = \prod_{i=1}^{H} \prod_{j=1}^{n_i} \left(\frac{1}{\sigma_i^2}\right)^{-1/2} \exp\left\{-\frac{(y_{ij} - x_i(t_j|\boldsymbol{\theta}))^2}{2\sigma_i^2}\right\}. \tag{26}$$

In real-world scenarios, while the DE systems depend on unknown parameters, researchers typically have prior knowledge about the parameters of the DE model. To incorporate the prior information into our estimation, we propose a Bayesian framework and treat the unknown parameters of the DE system as random variables. This Bayesian estimation approach enables us to find the statistical distribution of the unknown parameters based on the set of observed data, given prior information about the unknown parameters. The statistical distribution of parameters given the observed data is henceforth called the posterior distribution of the parameters $\pi(\boldsymbol{\Omega}|\mathbf{y})$ where $\boldsymbol{\Omega}$ denotes the set of all unknown parameters of the DE system. The posterior distribution then enables us to make statistical inferences about the uncertainty in the estimation procedure and quantify the characteristics of the DE system.

3.1 Sequential Monte Carlo

When we face DE system (25), it is reasonable to take into account the effect of space-time argument and the sequential process of the DE observations $\mathbf{y}_1, \ldots, \mathbf{y}_t$. This facilitates updating the posterior distribution of the unknown parameters sequentially as the DE variables are sequentially observed in the system. Sequential Monte Carlo (SMC), as a simulation-based approach, is a flexible technique in Bayesian statistics to sequentially estimate the posterior distribution. Due to the advent of cheap and powerful computing resources, the SMC appears a convenient tool to implement, in a parallel fashion, sequential computation of the posterior distribution in a general setting.

Let $\{\boldsymbol{\Omega}_t, t \in \mathbb{N}\}$ denote the Markov process showing the trajectory of unknown parameters over space-time with the prior distribution $\pi(\boldsymbol{\Omega}_0)$. The Markov property of the process indicates that the probability of process at space-time t only depends on the previous step of the state; that is

$$p(\boldsymbol{\Omega}_t|\boldsymbol{\Omega}_0, \ldots, \boldsymbol{\Omega}_{t-1}) = p(\boldsymbol{\Omega}_t|\boldsymbol{\Omega}_{t-1}),$$

where $p(\boldsymbol{\Omega}_t|\boldsymbol{\Omega}_{t-1})$ denotes the transition probability from $\boldsymbol{\Omega}_{t-1}$ to $\boldsymbol{\Omega}_t$ in the parameter space. Let $\{\mathbf{y}_t; t \in \mathbb{N}\}$ denote the sequence of the observations from the DE system where they are conditionally independent given the observed status of the parameter process with distribution $p(\mathbf{y}_t|\boldsymbol{\Omega}_t)$ for $t \geq 1$. For the sake of convenience in notations, let $\boldsymbol{\Omega}_{0:t}$ and $\mathbf{y}_{1:t}$ represent the parameter and DE observation sequences up to space-time t, respectively; That is, $\boldsymbol{\Omega}_{0:t} = (\boldsymbol{\Omega}_0, \ldots, \boldsymbol{\Omega}_t)$ and $\mathbf{y}_{1:t} = (\mathbf{y}_1, \ldots, \mathbf{y}_t)$.

In this section, our focus is to employ the SMC properties to obtain recursively the posterior distribution $\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t})$ at any space-time t. In order to do that, at space-time t, one can apply the Bayes rule and write

$$\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_{1:t}|\boldsymbol{\Omega}_{0:t})\pi(\boldsymbol{\Omega}_{0:t})}{\int p(\mathbf{y}_{1:t}|\boldsymbol{\Omega}_{0:t})\pi(\boldsymbol{\Omega}_{0:t})d\boldsymbol{\Omega}_{0:t}}. \tag{27}$$

From the joint distribution of $\boldsymbol{\Omega}_{0:t}$ and $\mathbf{y}_{1:t}$, one can also recursively update the posterior distribution $\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t})$ based on the posterior distribution of previous lags of the process by

$$\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t}) = \pi(\boldsymbol{\Omega}_{0:t-1}|\mathbf{y}_{1:t-1})\frac{p(\mathbf{y}_t|\boldsymbol{\Omega}_t)\pi(\boldsymbol{\Omega}_t|\boldsymbol{\Omega}_{t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}. \tag{28}$$

Accordingly, one can compute the posterior expectation of any characteristic of the DE system by

$$\mathbb{E}_{\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t})}[\mathbb{H}_t(\boldsymbol{\Omega}_{0:t})] = \int \mathbb{H}_t(\boldsymbol{\Omega}_{0:t})\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t})d\boldsymbol{\Omega}_{0:t} \tag{29}$$

Due to the complex structure of the uncertainty and multidimensional integration of the marginal distribution in the DE system, there is no analytical form for the posterior distributions (27) and (28); thus the conditional expectation (29) is not tractable too.

Importance sampling [51,52], as a practical solution to the intractability problem, uses an instrumental distribution to sample indirectly from the posterior distribution $\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t})$. Let $q(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t})$ represent an instrumental distribution whose domain includes the domain of the target posterior distribution (27). Accordingly, one can employ the importance sampling and rewrite (29) as

$$\mathbb{E}_{\pi(\boldsymbol{\Omega}_{0:t}|\mathbf{y}_{1:t})}[\mathbb{H}_t(\boldsymbol{\Omega}_{0:t})]$$

$$= \frac{\int \mathbb{H}_t(\mathbf{\Omega}_{0:t})\lambda(\mathbf{\Omega}_{0:t})q(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t})d\mathbf{\Omega}_{0:t}}{\int \lambda(\mathbf{\Omega}_{0:t})q(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t})d\mathbf{\Omega}_{0:t}}, \tag{30}$$

where the non-normalized importance weights $\lambda(\mathbf{\Omega}_{0:t})$ are given by

$$\lambda(\mathbf{\Omega}_{0:t}) = \frac{\pi(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t})}{q(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t})}. \tag{31}$$

Let $\{\mathbf{\Omega}_{0:t}^{(i)}\}$; $i = 1, \ldots, n$ represent n independent and identically distributed particles from the instrumental distribution $q(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t})$. Thus, using the importance sampling method, the Monte Carlo estimate of the quantity of interest $\mathbb{H}_t(\cdot)$ is given by

$$\widehat{\mathbb{H}_t(\mathbf{\Omega}_{0:t})} = \frac{\sum_{i=1}^{n} \mathbb{H}_t(\mathbf{\Omega}_{0:t}^{(i)})\lambda(\mathbf{\Omega}_{0:t}^{(i)})}{\sum_{i=1}^{n} \lambda(\mathbf{\Omega}_{0:t}^{(i)})} = \sum_{i=1}^{n} \mathbb{H}_t(\mathbf{\Omega}_{0:t}^{(i)})\Lambda(\mathbf{\Omega}_t^{(i)}) \tag{32}$$

where $\Lambda(\mathbf{\Omega}_{0:t}^{(i)})$, as normalized importance weights, are obtained by

$$\Lambda(\mathbf{\Omega}_t^{(i)}) = \frac{\lambda(\mathbf{\Omega}_{0:t}^{(i)})}{\sum_{i=1}^{n} \lambda(\mathbf{\Omega}_{0:t}^{(i)})}. \tag{33}$$

The importance sampling estimator (32), as a general framework of Monte Carlo, is convenient to implement. Despite this convenience, the iterative structure of the technique is not adequate to sequentially incorporate the new status of the process into estimating the $\pi(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t})$. As the new status of the sequence becomes available, one has to recompute all the important weights on the entire parameter space. This becomes computationally expensive for complex and nonlinear equations of motion.

Sequential Monte Carlo (SMC) [52,53], as a sequential architecture of importance sampling, can be considered as a solution to the iterative problem of general importance sampling. As observed from (31)–(33), in importance sampling, when a new status of the DE variable becomes available y_t, one has to re-compute the posterior distribution of the entire trajectory $\mathbf{\Omega}_{1:t}$ given the sequence $\mathbf{y}_{1:t}$. Hence, one requires re-computing even the importance weights of the previous states of the trajectory given the new status of the DE variable. Unlike the importance sampling method, the SMC sampler does not require re-computing the importance weights corresponding to the previous states of the trajectory $(\mathbf{\Omega}_0, \ldots, \mathbf{\Omega}_{t-1})$, when the new status of the DE variable becomes available. In other words, the importance weights of the previous states of the trajectory stay the same, and we no longer need to re-compute them. We only need to compute the posterior distribution of the most recent state of the

trajectory, given the new data. Treating the instrumental distribution based on the previous states as the full marginal distribution, the instrumental distribution is thus updated by

$$q(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t}) = q(\mathbf{\Omega}_{0:t-1}|\mathbf{y}_{1:t-1})q(\mathbf{\Omega}_t|\mathbf{\Omega}_{0:t-1}, \mathbf{y}_{1:t}). \tag{34}$$

In a similar fashion as (34), one can recursively show that

$$q(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t}) = \pi(\mathbf{\Omega}_0)\prod_{k=1}^{t} q(\mathbf{\Omega}_k|\mathbf{\Omega}_{0:k-1}, \mathbf{y}_{1:k}). \tag{35}$$

From the sequential representation (35), one can easily show that the non-normalized importance weights (33) can be sequentially updated by

$$\lambda(\mathbf{\Omega}_{0:t}^{(i)}) = \lambda(\mathbf{\Omega}_{0:t-1}^{(i)})\frac{p(\mathbf{y}_t|\mathbf{\Omega}_t^{(i)})\pi(\mathbf{\Omega}_t^{(i)}|\mathbf{\Omega}_{0:t-1}^{(i)})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})\pi(\mathbf{\Omega}_t^{(i)}|\mathbf{\Omega}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})}, \tag{36}$$

and consequently, the normalized importance weights are sequentially updated by

$$\Lambda(\mathbf{\Omega}_t^{(i)}) \propto \Lambda(\mathbf{\Omega}_{t-1}^{(i)})\frac{p(\mathbf{y}_t|\mathbf{\Omega}_t^{(i)})\pi(\mathbf{\Omega}_t^{(i)}|\mathbf{\Omega}_{0:t-1}^{(i)})}{\pi(\mathbf{\Omega}_t^{(i)}|\mathbf{\Omega}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})}. \tag{37}$$

As a special case of (35), one can employ the prior distribution in the SMC framework [52,53]. In this case, the instrumental distribution is given by

$$q(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t}) = \pi(\mathbf{\Omega}_0)\prod_{k=1}^{t} \pi(\mathbf{\Omega}_k|\mathbf{\Omega}_{k-1}). \tag{38}$$

In this case, from (38) and the fact that $p(\mathbf{y}_t|\mathbf{y}_{1:t-1})$ is constant over the entire trajectory of $\mathbf{\Omega}_{0:t}$, one can easily update the importance weights of the i-th particle by

$$\Lambda(\mathbf{\Omega}_t^{(i)}) \propto \Lambda(\mathbf{\Omega}_{t-1}^{(i)})p(\mathbf{y}_t|\mathbf{\Omega}_t^{(i)}). \tag{39}$$

Although the SMC technique is well suited to accommodate sequentially the new data into estimation, the posterior distribution of the particles very quickly becomes skewed after only a few steps such that only a few particles will have a non-zero probability [53,54]. Consequently, the Monte Carlo chain will not be able to sample from all aspects of target $\pi(\mathbf{\Omega}_{0:t}|\mathbf{y}_{1:t})$. To deal with this degeneracy issue, a re-sampling is added to SMC to eliminate sequentially particles with low importance weights and at the same time augment the particles from high-density areas. To implement the re-sampling step, one can take n random draws with replacement from the collection of the particles $\{\mathbf{\Omega}_{0:t}^{(1)}, \ldots, \mathbf{\Omega}_{0:t}^{(n)}\}$ with probabilities corresponding to the important weights

$\{\Lambda(\boldsymbol{\Omega}_t^{(1)}), \ldots, \Lambda(\boldsymbol{\Omega}_t^{(n)})\}$. Let $n_t^{(i)}$ denote the number of off-springs from the particle $\boldsymbol{\Omega}_{0:t}^{(i)}, i = 1, \ldots, n$. It is easy to see that the particle survives and contributes to the posterior distribution when $n_t^{(i)} > 0$; otherwise, the particle dies. The important sampling and the re-sampling steps are finally alternated to update the important weights and filter the particles sequentially to obtain the samples from the posterior distribution.

3.2 Cross-validated neural networks

Neural networks (NNs), inspired by the human neural system, comprise a network of connected neurons. These connections enable the neurons to send information from one layer to another. According to the flexibility and power of the NNs, they have been increasingly exploited in recent years as a reliable predictive model for solving differential equations. The power of the NNs enables us to reformulate finding solutions to the DE system to a parametric estimation of the DE variables by minimizing the prediction errors. NNs consist of a multi-layer perceptron whose layers include neurons connecting the layers of the network to each other. These connections enable the NN to estimate the functional form of the DE variables. In this subsection, we describe how NNs use the parameter estimates developed by the SMC method, from Sect. 3.1, to predict the complex and nonlinear forms of the DE variables in the hyperbolic class of 4d.

Let $\mathcal{N}(\mathbf{x}(t|\boldsymbol{\Omega}), t, \boldsymbol{\phi})$ denote the NN estimate, consisting of L hidden layers, for the DE variable $\mathbf{x}(t)$ form DE system (25). Each neuron of the NN is connected with another in the next layer via a linear regression model

$$\begin{cases} \mathbf{Z}_j^l = \mathbf{W}^l \text{NN}^{l-1}(\mathbf{x}(t), t, \boldsymbol{\phi}) + \mathbf{b}^l, \\ \text{NN}^l(\mathbf{x}(t), t, \boldsymbol{\phi}) = a(\mathbf{Z}_j^l), \qquad l = 1, \ldots, L. \end{cases} \tag{40}$$

such that $\text{NN}^{l-1}(\cdot)$ represents the response observed from the l-th layer, $\boldsymbol{\phi} = (\mathbf{W}^1, \ldots, \mathbf{W}^L, \mathbf{b}^1, \ldots, \mathbf{b}^L)$ represents the set of all unknown parameters, a denotes a non-linear activation function, $\mathbf{W}^l$ and $\mathbf{b}^l$ show the weight matrix and bias vector of the l-th layer, respectively [55,56]. Finally, the NN estimate of the DE variables $\mathbf{x}(t)$ are obtained as a solution to the squared loss function

$$\widehat{\mathcal{N}}(\mathbf{x}(t|\boldsymbol{\Omega}), t, \boldsymbol{\phi}) = \arg\min_{\boldsymbol{\Omega}} \left(\mathcal{N}(\mathbf{x}(t|\boldsymbol{\Omega}), t, \boldsymbol{\phi}) - \mathbf{x}(t) \right)^2. \tag{41}$$

To handle the optimization (41), the NNs implement a series of forward and backward propagation steps to find the final solution to the DE system (25). For more details about the theory and applications of the NNs, readers are referred to [55–57] and references therein.

In this paper, we first plan to find the Bayesian estimate for the parameters of the equations of motion. The Bayesian proposals are then stacked into the equations of motion to find the Bayesian stacked NN solvers. To this end, as described in Sect. 3.1, we use the SMC method and find the posterior distribution $\pi(\boldsymbol{\Omega}|\mathbf{y})$. Let $\boldsymbol{\Omega}_1^*, \ldots, \boldsymbol{\Omega}_M^*$ denotes M Bayesian SMC proposals for the parameters of the DE system. Given the posterior candidate $\boldsymbol{\Omega}_m^*$, let $\widehat{\mathcal{N}}(\mathbf{x}(t|\boldsymbol{\Omega}_m^*), t, \boldsymbol{\phi})$ denote the NN-based estimate of the DE variable $\mathbf{x}(t)$ for $m = 1, \ldots, M$. From a probabilistic perspective, the estimate $\widehat{\mathcal{N}}(\mathbf{x}(t|\boldsymbol{\Omega}_m^*), t, \boldsymbol{\phi})$ will be the true solution to the DE system (25) with probability $\pi(\boldsymbol{\Omega}_m^*|\mathbf{y})$ for $m = 1, \ldots, M$.

Recently, Hatefi et al. [45] proposed the Bayesian model averaging to stack the NN-based estimates in predicting the critical solution for the elliptic class of 4d. Following [45], using the SMC candidates $\boldsymbol{\Omega}_1^*, \ldots, \boldsymbol{\Omega}_M^*$ and the training set $\mathbf{y}$ of size n, the posterior distribution of the $\widehat{\mathcal{N}}(\mathbf{x}(t|\boldsymbol{\Omega}_m^*), t, \boldsymbol{\phi})$ at a fixed space-time t, is given by

$$\mathbb{P}(\widehat{\mathcal{N}}(\mathbf{x}(t|\boldsymbol{\Omega}), t, \boldsymbol{\phi})|\mathbf{y}) = \sum_{m=1}^M \mathbb{P}(\widehat{\mathcal{N}}(\mathbf{x}(t|\boldsymbol{\Omega}_m^*), t, \boldsymbol{\phi})|\mathbf{y})\pi(\boldsymbol{\Omega}_m^*|\mathbf{y}). \tag{42}$$

Using (42), Hatefi et al. [45] proposed the mean posterior of the NN-based estimates to stack the NN candidates in predicting the DE variables in the elliptic class of equations of motion. Despite the simplicity of the posterior mean, when the posterior distribution has multiple high-density areas, the posterior mean may not be able to capture the different solutions of the DE system. To handle the problem, we propose the idea of Leave-One-Out Cross-validation (LOOCV) to combine information from all high-density areas of the posterior and develop cross-validated NN-based estimates to more accurately estimate the DE variables under various solutions of the hyperbolic class equations of motion.

Let $\widehat{\mathcal{N}}(\mathbf{x}|\boldsymbol{\Omega}^*) = \left[\widehat{\mathcal{N}}_1(\mathbf{x}(t|\boldsymbol{\Omega}_1^*), t, \boldsymbol{\phi}), \ldots, \widehat{\mathcal{N}}_M(\mathbf{x}(t|\boldsymbol{\Omega}_M^*), t, \boldsymbol{\phi})\right]^\top$ represent the design matrix of the M estimates, corresponding to SMC candidates. As $\widehat{\mathcal{N}}_1(\mathbf{x}(t|\boldsymbol{\Omega}_1^*), t, \boldsymbol{\phi}), \ldots, \widehat{\mathcal{N}}_M(\mathbf{x}(t|\boldsymbol{\Omega}_M^*), t, \boldsymbol{\phi})$ are M estimates of the DE variable $\mathbf{x}(t)$ at space-time t, the estimates may be linearly dependent. This arises the co-linearity problem in the NN-based design matrix $\widehat{\mathcal{N}}(\mathbf{x}|\boldsymbol{\Omega}^*)$. To this end, under squared error loss with l_2 penalty, we can stack the NN-based estimates using the penalized linear regression model [55]. One can estimate the coefficients of the regression model $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_M)$ by

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\beta} \left[\left(\mathbf{y} - \widehat{\mathcal{N}}(\mathbf{x}|\boldsymbol{\Omega}^*)\right)^\top \left(\mathbf{y} - \widehat{\mathcal{N}}(\mathbf{x}|\boldsymbol{\Omega}^*)\right) + \lambda\boldsymbol{\beta}^\top\boldsymbol{\beta} \right]. \tag{43}$$

According to the properties of the least square under l_2 penalty, one can easily show that the solution to (43) is given by

$$\widehat{\beta} = \left[\widehat{\mathcal{N}}(\mathbf{x}|\boldsymbol{\Omega}^*)^\top \widehat{\mathcal{N}}(\mathbf{x}|\boldsymbol{\Omega}^*) + \lambda \mathbb{I}\right]^{-1} \widehat{\mathcal{N}}(\mathbf{x}|\boldsymbol{\Omega}^*)\mathbf{y} \qquad (44)$$

When the posterior distribution has multiple modes, the least square estimate (44) may assign unfair weights to some complex NN candidates based on the training set. To deal with this problem, we develop the LOOCV estimate of the weights where leaving one observation out in each iterative training step to find the best coefficient estimates. Let $\widehat{\mathcal{N}}_m^{-i}(\mathbf{x}(t|\boldsymbol{\Omega}_m^*), t, \boldsymbol{\phi})$ denote the NN-based estimate for the DE variables at space-time t using the m-th SMC proposal when the i-th observation in the training set has been removed. From (43), the LOOCV estimate of the weights are given by

$$\widehat{\beta}_{loocv} = \arg\min_{\beta}$$
$$\left[\sum_{i=1}^{M} \left(\mathbf{y} - \widehat{\mathcal{N}}^{-i}(\mathbf{x}|\boldsymbol{\Omega}^*)\right)^\top \left(\mathbf{y} - \widehat{\mathcal{N}}^{-i}(\mathbf{x}|\boldsymbol{\Omega}^*)\right) + \lambda \boldsymbol{\beta}^\top \boldsymbol{\beta}\right].$$
$$(45)$$

From (45), the LOOCV-based NN estimate of the DE variables at space-time t, using the SMC estimates $\boldsymbol{\Omega}_1^*, \ldots, \boldsymbol{\Omega}_1^*$ is given by

$$\widehat{\mathbf{x}}(t) = \sum_{m=1}^{M} \widehat{\beta}_{loocv} \, \widehat{\mathcal{N}}_m(\mathbf{x}(t|\boldsymbol{\Omega}_m^*), t, \boldsymbol{\phi}) \qquad (46)$$

The LOOCV-based NN estimates (46) cross-validates iteratively the NN candidates on $\widehat{\mathcal{N}}^{-i}(\mathbf{x}|\boldsymbol{\Omega}^*)$; thus the final estimates avoid giving unfair weights to spurious SMC proposals for the training set. It takes better into account the multiple high-density areas of the posterior distribution and consequently, it is expected to more efficiently estimate the DE variables of the system (25), giving the SMC estimates $\boldsymbol{\Omega}_1^*, \ldots, \boldsymbol{\Omega}_1^*$.

4 Numerical studies

The equations of motion for hyperbolic class in four dimensions have five singular points; however, the relevant range for z, that contains the proper information lies between the two singularities

$$z = 0, \qquad (47)$$
$$z = z_+, \quad b(z_+) = z_+. \qquad (48)$$

In particular, $z = z_+$ is an event horizon, which is the homothetic horizon. Thus, it is a coordinate singularity, and τ must be regular across it, which is equivalent to the finiteness of $f''(z)$ as $z \to z_+$. In fact the vanishing of the divergent part

of $f''(z)$ leads to a complex-valued constraint at z_+ as

$$\mathbb{C} \ni G(b(z_+), f(z_+), f'(z_+)). \qquad (49)$$

From time scaling, the regularity of τ, and the residual symmetries in the equations of motions, one can show the initial boundary conditions as

$$b(0) = 1, \, f(0) = 1 + i x_0, x_0 > 0, \quad f'(0) = 0, \qquad (50)$$

where x_0 is a real parameter as the initial value of the equations of motion. Thus, the system results in one parameter ω and two constraints which are the real and imaginary parts of G. Therefore, the system is handled by discrete solutions where the CSS solutions are numerically investigated.

Here we plan to estimate the self-similar solutions using the Bayesian framework for the hyperbolic class in 4d. In this framework, we treat the parameter of the equations of motion ω as a random variable and then use the SMC approach to find the posterior distribution. The posterior distribution allows us to take into account the numerical measurement errors in estimating the critical collapse functions of the the system.

The equations of motion of the axion-dilaton system have already been studied in various dimensions and also for different ansatz [23,44]. Hatefi et al. [42,43] applied the statistical regression models using Fourier-based and spline smothers to estimate the critical collapse functions. In a recent publication, Hatefi et al. [44] constructed a solver based on NNs and showed there was no solution in higher dimensions for the parabolic class of black holes. Particularly [23] employed a root-finding method to numerically determine all the parameters of the equations of motion.

Due to the highly nonlinear equations of motion, researchers typically have to use a series of numerical approaches to simplify the equations and also to keep track of the parameters of the models. For instance, these techniques include imposing non-trivial constraints on the equations of motions such as finiteness of $f''(z)$ as $z \to z_+$. One might also point out one more constraint, namely, the vanishing of the divergent part of $f''(z)$ that generates the complex-valued constraint at z_+. This implies that the real and imaginary parts of G must vanish at $z \to z_+$. The equations in the hyperbolic case are not solvable analytically, hence [39] used the profile-root finding method and applied the discrete optimization on the coordinates of the extended parameter space of the equations. More specifically, the equations are approximated by the first two orders of the Taylor expansions to make the root-finding step for the parameters of the equations tractable. Next, [39] could clearly estimate the parameters of the equations of motion by applying a grid search optimization on the coordinates in the system. After discarding the spurious roots in the domain, [39] found out the solution to the equations of motion and then could estimate the critical collapse functions.

Although [23,42–44] explored the hyperbolic class of equations motion in deterministic approach, in this research

through a stochastic perspective, we construct a Bayesian method using SMC approach to assess the measurement errors into the estimation of the critical functions. Recently Hatefi et al. [45] proposed Bayesian mean using Hamiltonian Monte Carlo to find solutions to the equations of motion in the elliptic class of 4d. Unlike [45], the hyperbolic equations of motion in 4d result in multiple solutions where the critical collapse functions have overlap domains under the multiple solutions. To deal with the complexity of the measurement errors in the hyperbolic equations of motion, for the first time in the literature on the axion-dilaton system, we proposed the Sequential Monte Carlo approach to derive the posterior distributions of the parameters. Unlike [45] where they proposed Bayesian model averaging to stack the NN-based estimates, here we propose the l_2 penalized Leave-One-Out cross-validation to stack the NN-based estimates for critical collapse functions. The approach enables us to assign the optimum weights to NN-based estimates and hence better estimate the critical functions corresponding to multiple solutions of the equations of motion.

The SMC Bayesian estimates provide information for all possible parameters' outcomes, that may take place in the numerical experiments, from the posterior distribution. They also allow researchers to embed this complexity in estimating critical functions. In this numerical study, we consider equations of motion (20) as the DE system of interest. The system leads to three critical collapse functions $b_0(z)$, $|f(z)|$ and $\arg(f(z))$ where we treated these functions as the DE variables of the system that must be estimated. Since all the DE variables of system (20) are numerically and simultaneously solved, it makes sense to assume that the observed DE variables also have the same standard deviation σ parameter which actually represents the variability in the numerical experiments. Hence, $\mathbf{\Omega} = (\omega, \sigma)$ encodes the set of all unknown parameters of the model. We used Python Package Pymc3 [58] to implement the SMC approach and find the posterior distribution of $\mathbf{\Omega}$. In order to investigate the effect of the prior information on the likelihood parameters in predicting of our critical functions, we assigned two different prior distributions for ω. The prior distributions include non-informative uniform distribution between [0.3, 1.5]. We assign the second prior distribution to be the Gaussian distribution with a mean of 1.20 and a standard deviation of 0.2. We also consider the half-Cauchy distribution with scale parameter 0.5 as the prior distribution for parameter σ to capture the uncertainty involved in the likelihood function.

We show the posterior distributions of parameters ω and σ in Figs. 1 and 6 under Gaussian and Uniform prior distributions, respectively. In each figure, we show two independent realizations of the SMC chain for the posterior distributions and their corresponding trace plots. It is clear that the posterior distributions of ω are multi-modal and have multiple high-density areas. Interestingly, this finding is com-

patible with the literature where [39] shows there must be at least three solutions to hyperbolic equations of motion in 4d. Under both prior distributions, we observe that there is a dominant high-density area almost ranging between [1.25, 1.4]. This corresponds to the α-solution of [39]. In addition, we also observe that both figures confirm there is a small high density on the left tail of the posterior distribution of ω which is interestingly compatible with the β-solution to the equations of motion in [39]. It should be noted that for example under Uniform prior, the probability that the deterministic α-solution of [39] be the true solution to the hyperbolic equations motion is almost 10%; that is $\pi(\omega \in 1.36 \pm 10^{-2}) \propto 0.10$. Moreover, the probability that the true solution appears to be smaller than 1.10 is given by $\pi(\omega \le 1.10) \propto 0.05$.

As described in Sect. 3.2, we construct the NN-based estimates of the critical collapse functions using the SMC proposals from the posterior distribution $\pi(\omega|\mathbf{y})$. To do that, we take $L = 200$ samples $\omega_1^*, \ldots, \omega_L^*$ from $\pi(\omega|\mathbf{y})$. Treating the posterior candidates ω_l^*, $l = 1, \ldots, L$ as the true value of the parameter in the equations of motion, we applied fully connected NNs to solve the DE system and find the NN-based estimates of the critical collapse functions. To do so, we used Python Package NeuroDiffEq [59] to carry out the neural networks for differential equations with 4 hidden layers where each layer consists of 16 neurons. We then ran the NNs for 1000 epochs and estimates the critical collapse function at 1000 equally spaced space-time points $z_i \in [0, 1.44]$ for $i = 1, \ldots, 1000$. We finally obtained $L = 200$ NN-based estimates for the critical collapse functions corresponding to L realizations from the posterior distributions. From a probabilistic perspective, each of these L realizations of NN-based estimates can occur to be the true form of the critical functions with probabilities $\pi(\omega_l^*|\mathbf{y})$ for $l = 1, \ldots, L$. To better represent this probabilistic perspective and sampling variability from $\pi(\omega_l^*|\mathbf{y})$, we show ten randomly selected realizations of the posterior NN-based estimates for the critical function in Figs. 2 and 7 under Gaussian and Uniform priors. We see that two different patterns are almost observed for critical collapse functions on the same domain which is compatible with the multiple solutions available in the literature for the equations of motion under the hyperbolic class of 4d.

In the next step of the numerical study, we plan to stack the $L = 200$ NN-based candidates in estimating the critical collapse functions. Using a root-finding based on general relativity, [39] showed that there are three solutions in the domain of the equations of motion in the hyperbolic class of 4d. These solutions correspond to $\omega_1 = 1.362$ (α-solution), $\omega_2 = 1.003$ (β-solution) and $\omega_3 = 0.005$ (γ-solution). It was discussed in [39] that the γ-solution is not a stable solution due to the fact that $\mathrm{Im} f(0)$ is so small also the z_+ root-finding gets affected by numerical noise ($\omega = 0.541$, $v(0) = 0.0059$, $z_+ = 8.44$), and hence the
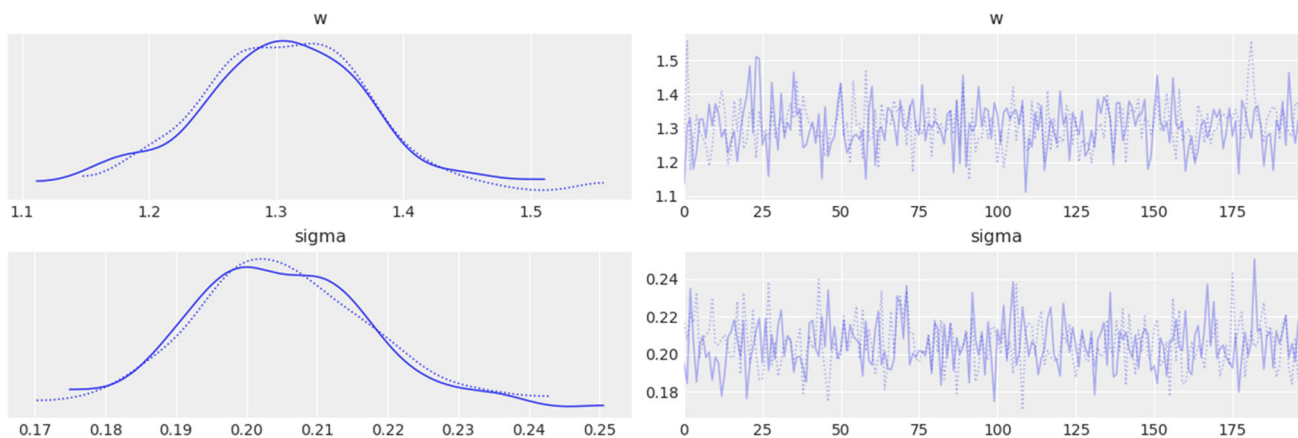
Fig. 1 The posterior distributions and their trace-plots of two SMC chains (shown by solid and dotted blue lines) for parameters ω (w) and σ (sigma) under Gaussian prior distribution
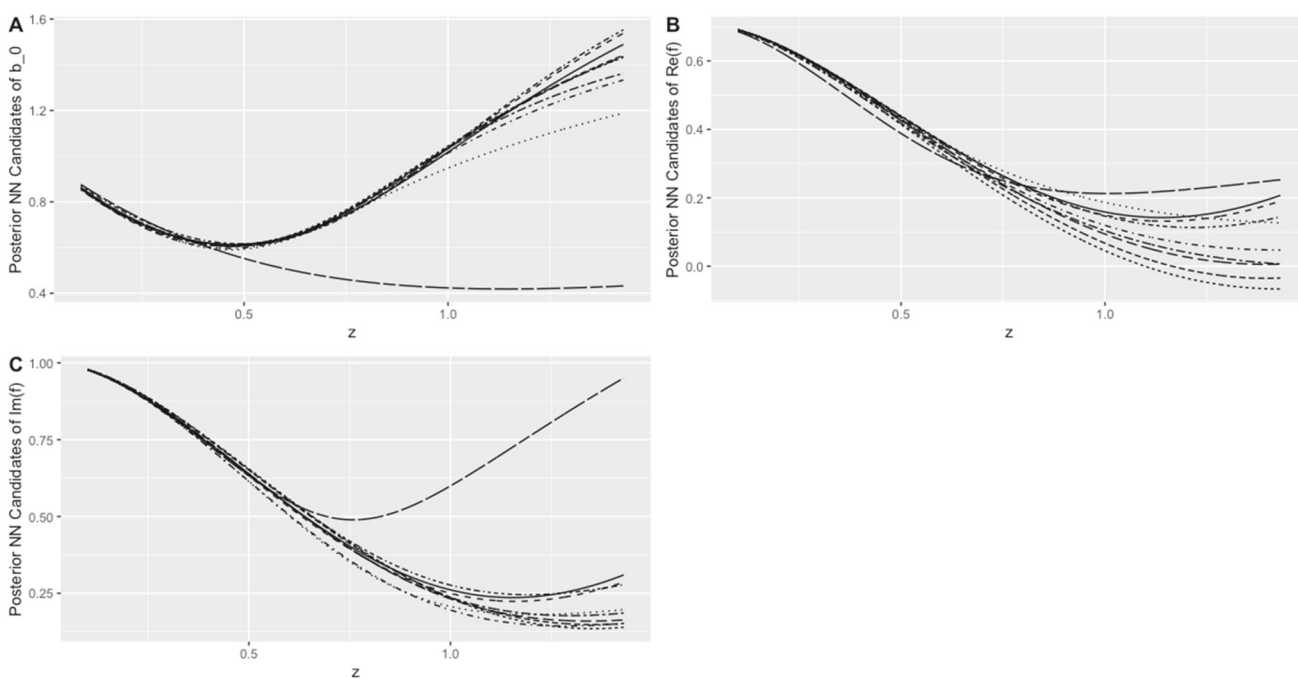

Fig. 2 Ten randomly selected NN-based estimates for DE variables $b_0(z)$ (**A**), $\mathrm{Re}(f(z))$ (**B**) and $\mathrm{Im}(f(z))$ (**C**) corresponding to ten SMC samples from the support of the posterior distribution $\pi(\omega|\mathbf{y})$ where Gaussian distribution was used as prior distribution. We show each estimate with a different line type

quality is not perfect and so it may be a spurious solution due to numerical noise. For this reason, in this research, we focused on the two α and β solutions as two available solutions to the equations of motion in the literature. The critical collapse functions corresponding to α-solution range between [0, 1.44], while the functions range between [0, 3.29] corresponding to β-solution. In order to investigate the performance of developed models in estimating the functional form of the critical collapse functions under both solutions, we focus on the common areas between the two

scenarios and ran investigated the NN-based estimates in [0,1.44].

Following Hatefi et al. [45], we first applied the Bayesian model averaging method and computed the posterior mean of the L NN-based estimates at each space-time point z_i, $i = 1, \ldots, 1000$. Henceforth this estimate is called the mean stacked NN-base estimate of the critical collapse functions. We also computed the NN-based estimates for critical collapse functions under $\omega = 1.362$ and $\omega = 1.003$ corresponding to estimates under α and β solutions. We treated these fixed NN-based estimates as the true forms of the critical col-
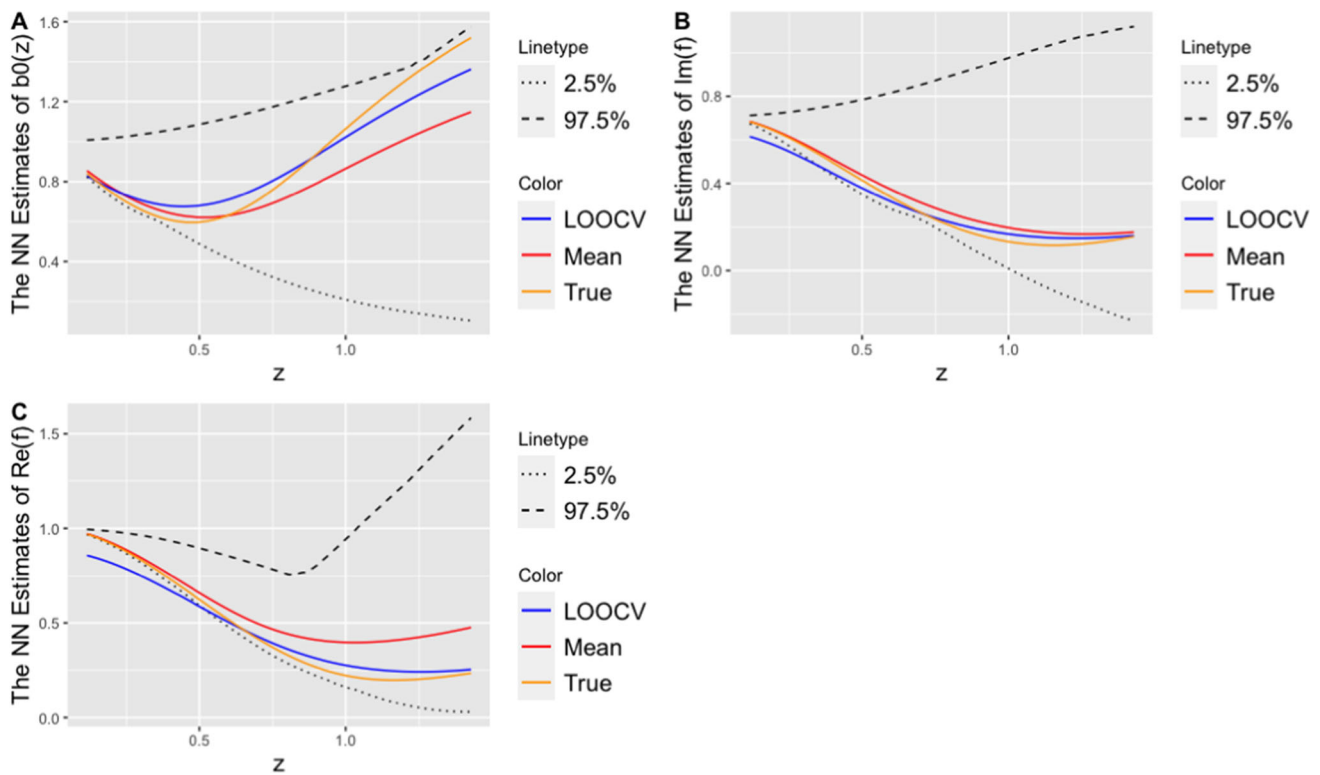
Fig. 3 The posterior mean (red), LOOCV (blue) and the true, using $\omega = 1.362$, (orange) NN-based estimates of the critical collapse functions $b_0(z)$ (**A**), Im($f(z)$) (**B**) and Re($f(z)$) (**C**) corresponding to the α-solution scenario of the hyperbolic equations of motion. The dotted and dashed lines show, respectively, the lower and upper bounds of the 95% Bayesian credible intervals under Gaussian prior

lapse function under two different solutions to the equations of motion. We then applied the LOOCV technique to stack the NN candidates in estimating the critical functions when we used 100 random observations from the true NN-based estimates as the taring sets for the LOOCV-based NN methods. On the other side, we also obtained the 95% Bayesian credible intervals for the critical collapse functions. To do that, we computed the 2.5 and 97.5 percentiles of the NN-based estimates, respectively, as the lower and upper bounds of the interval at each space-time point.

Figures 3 and 8 show the performance of the posterior mean NN-based estimates, the true NN-based estimates, the LOOCV NN-based estimates as well as the 95% credible intervals in estimating the critical collapse functions corresponding to α-solution under Gaussian and Uniform priors, respectively. Also Figs. 4 and 9 show the results of their counterpart NN-based estimates corresponding to β-solution under Gaussian and Uniform prior distributions, respectively. Since the posterior mean NN-based estimates and 95% credible intervals aggregate the $L = 200$ posterior NN-based estimates regardless of their corresponding population, these two methods remain robust and proposed the same estimates in predicting the critical collapse functions under both α- and β- solutions to the equations of motion. Unlike poste-

rior mean proposals, we recommend the LOOCV-based estimates if one is interested in aggregating the NN candidates to more accurately estimate the form and curvature of the critical collapse functions corresponding to specific solutions. We see that the LOOCV-based method, on average, more accurately estimates the critical functions for both α and β solutions, because the method stacks the $L = 200$ Bayesian NN candidates by developing the best linear combination of the posterior NN candidates.

Last but not least, we investigated the convergence of the NN-based using SMC samples of ω under both Gaussian and Uniform prior distributions. To do that, we computed the difference between the training and test loss values in the last epoch of the NN-based solver for all the $L = 200$ SMC samples $\omega_l^*, l = 1, \ldots, 1000$ from the posterior distributions. Figures 5 and 10 report the trace and box plots of the loss differences when the Gaussian and Uniform distributions were, respectively, used as the prior distributions. From Figs. 5 and 10, we observe that loss differences are almost less then 1.5×10^{-2}, fluctuating on average around zero for all the $L = 200$ realizations. This confirms the convergence of the NN-based solvers using SMC samples after 1000 epochs even in the hyperbolic equations of motion of four dimen-
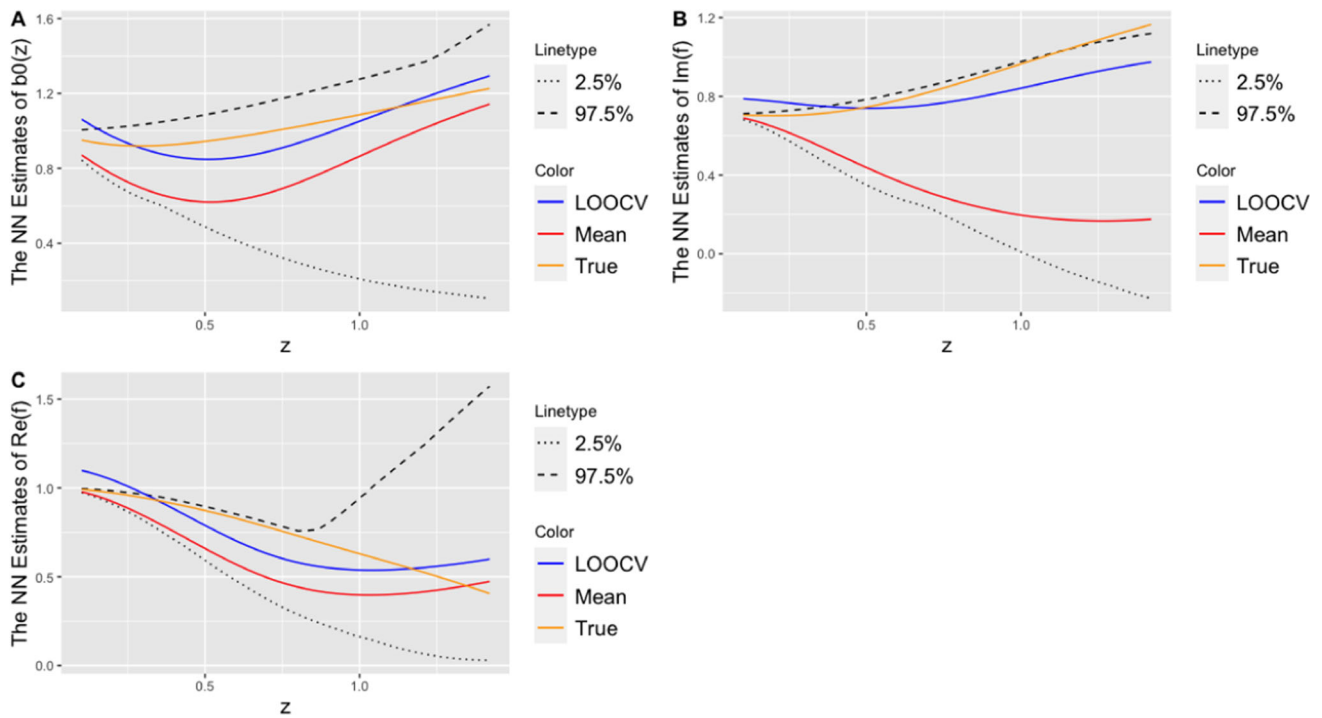
Fig. 4 The posterior mean (red), LOOCV (blue) and the true, using $\omega = 1.003$, (orange) NN-based estimates of the critical collapse functions $b_0(z)$ (**A**), Im($f(z)$) (**B**) and Re($f(z)$) (**C**) corresponding to the β-solution scenario of the hyperbolic equations of motion. The dotted and dashed lines show, respectively, the lower and upper bounds of the 95% Bayesian credible intervals under Gaussian prior
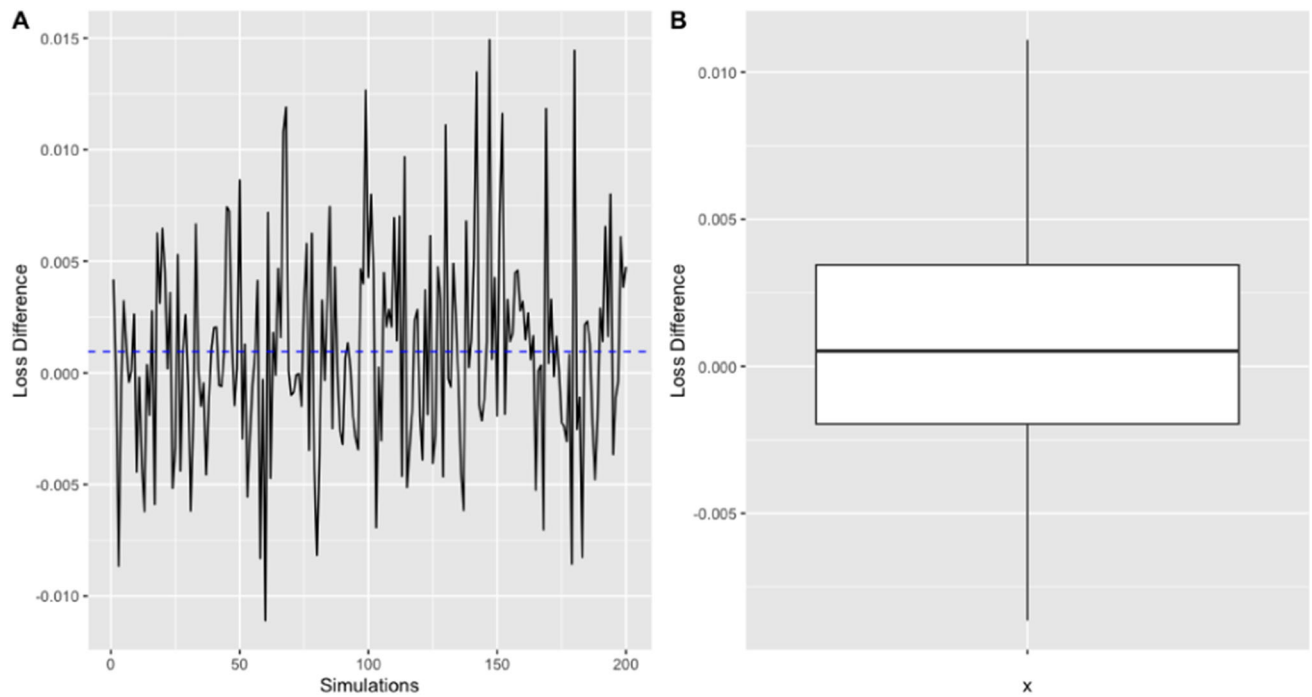
Fig. 5 The trace- and box- plots of the differences between train and test loss values in the NN-based solvers using $L = 200$ SMC samples from the posterior distribution of $\pi(\omega|\mathbf{y})$ when Gaussian distribution was used as prior

sions where the posterior distribution of the system parameter appear multi-modal with multiple high-density areas.

Finally, it is important to highlight that if instead of taking continuous self-similarity, one only assumes discrete self-similarity, then the discrete scale transformation is compensated just by an element of $SL(2, Z)$ in the set $SL(2, R)$ transformation. It would also be very interesting to discover the fact that how all the critical exponents would depend on the modular transformations as well.

5 Conclusions

This paper proposes a new formalism based on Sequential Monte Carlo (SMC) and artificial neural networks (NNs) to model the hyperbolic class of the spherical gravitational self-similar solutions in four dimensions. Due to the nature of highly non-linear ordinary differential equations for the axion-dilaton configurations, in the literature, researchers typically have to employ various constraints as well as different numerical approximation methods to keep track of equations. For instance, in hyperbolic equations of motion, [39] had to apply the constraints, including the finiteness of $f''(z)$ as $z \to z_+$ and the vanishing of the divergent part of $f''(z)$ which generates a complex-valued constraint at z_+. They also employed numerical grid search discrete optimization methods on the extended coordinates of the equations to eliminate the spurious roots and estimate the equations' parameters and the self-similar black hole solutions. Due to the sophisticated form of the equations and the vital role of the parameters, researchers usually have to overlook the measurement errors imposed in exploring the parameters through various numerical methods.

Here, we propose a new method to incorporate the measurement errors, involved in parameter estimation, into our statistical models in exploring the solutions to the equations of motion. Recently Hatefi et al. [45] applied the Hamiltonian Monte Carlo method to find solutions to the equations of motion in the elliptic class of four dimension in a Bayesian framework. Unlike [45], the hyperbolic equations of motion in four dimensions suffer from multiple solutions where the critical collapse functions have overlap domains under these solutions. To deal with this challenge, for the first time in the literature on the axion-dilaton system, we proposed the SMC approach to derive the posterior distribution of the parameters. The posterior distribution reveals all the possible solutions in estimating the parameter of the equations of motion. It is also important to highlight that the posterior distribution confirms the deterministic α and β solutions found in the literature for the hyperbolic class in four dimensions. Unlike methods in the literature, in this paper, we proposed the l_2 penalized Leave-One-Out Cross-validation (LOOCV) to optimally combine the Bayesian NNs

candidates. The approach enables us to determine the optimum weights while dealing with the co-linearity issue in the NN-based estimates and better predict the critical functions corresponding to multiple solutions of the hyperbolic equations of motion.

Using SMC samples from the posterior distribution, we then developed NN estimates based on the posterior mean and LOOCV as well as the 95% credible intervals in estimating the critical collapse functions corresponding to multiple solutions of the equations of motion. Because the posterior mean NN-based estimates and 95% credible intervals aggregate all the posterior NN candidates regardless of the multiple solutions of the system, these two methods remain robust and proposed the same estimates in predicting the critical collapse functions under both α- and β- solutions. Unlike the posterior mean proposal, we recommend the LOOCV-based estimates if one is interested in the optimum linear combination of the posterior NN candidates to improve the estimating of the critical collapse functions corresponding to established α- and β-solutions in hyperbolic equations of motions in four dimensions.

Now if we compare the Bayesian method with the realistic NN approach, then one gets to know that the developed Bayesian credible intervals actually contain the definite estimate as one possible candidate in the estimation of the critical collapse functions. Indeed, Unlike the estimation of [44], the Bayesian approach remains concrete against measurement errors in estimating ω due to the fact that all these Bayesian estimations have already had all the possibilities of the parameter for the domain of the posterior distribution. From a physical point of view, our results clarify that the universality of the Choptuik phenomena [17] is not satisfied. There may exist some universal behaviour which might be hidden in combining the critical exponents and other parameters of the theory. Nevertheless, our efforts provide some clear evidence that one cannot expect to transfer the standard expectations of Statistical Mechanics to the critical gravitational phenomena.
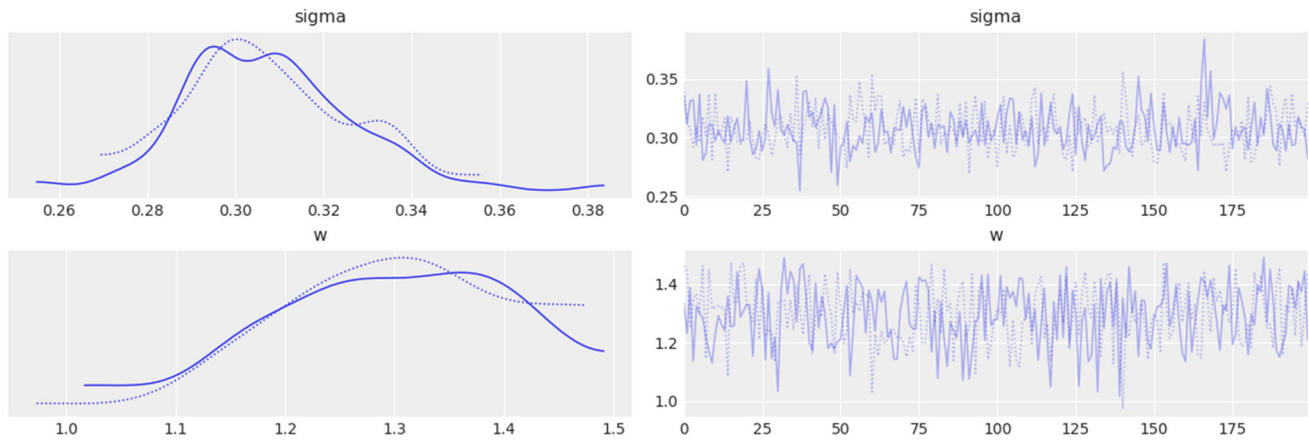
Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment:Data will be made available on request.]

Appendix

See Figs. 6, 7, 8, 9 and 10.

Fig. 6 The posterior distributions and their trace-plots of two SMC chains (shown by solid and dotted blue lines) for parameters ω (w) and σ (sigma) under Uniform prior distribution
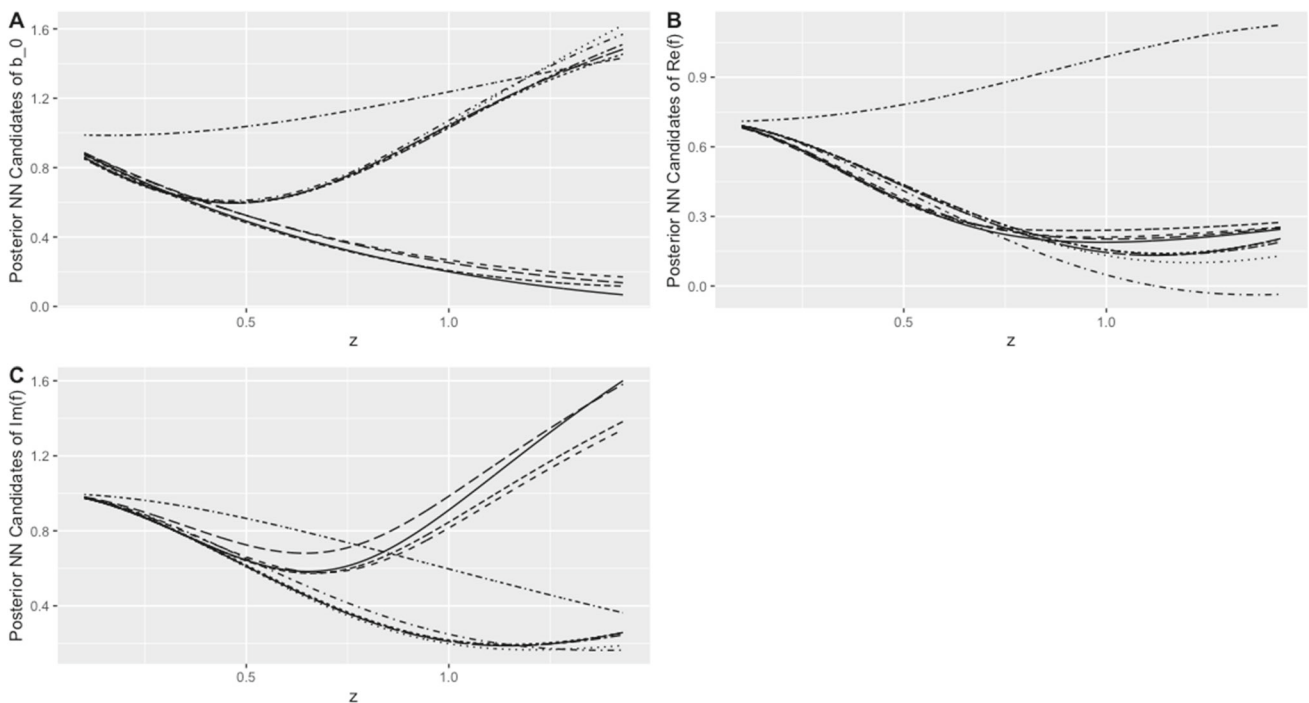
**Fig. 7** Ten randomly selected NN-based estimates for DE variables $b_0(z)$ (**A**), $\mathrm{Re}(f(z))$ (**B**) and $\mathrm{Im}(f(z))$ (**C**) corresponding to ten SMC samples from the support of the posterior distribution $\pi(\omega|\mathbf{y})$ where Uniform distribution was used as prior distribution. We show each estimate with a different line type
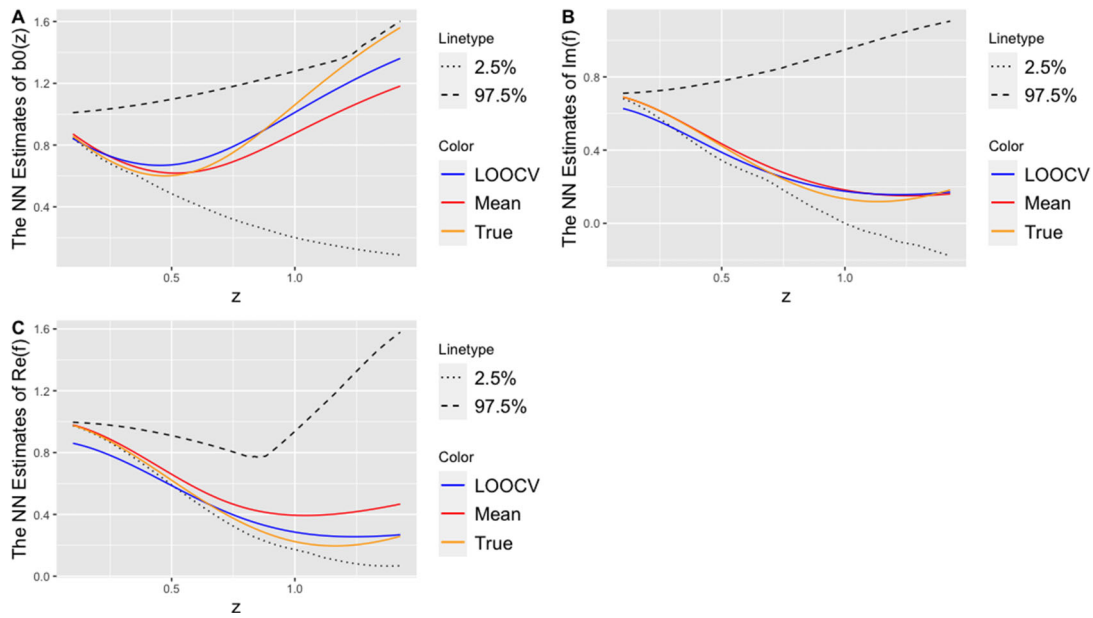
Fig. 8 The posterior mean (red), LOOCV (blue) and the true, using $\omega = 1.362$, (orange) NN-based estimates of the critical collapse functions $b_0(z)$ (**A**), $\text{Im}(f(z))$ (**B**) and $\text{Re}(f(z))$ (**C**) corresponding to the α-solution scenario of the hyperbolic equations of motion. The dotted and dashed lines show, respectively, the lower and upper bounds of the 95% Bayesian credible intervals under Uniform prior
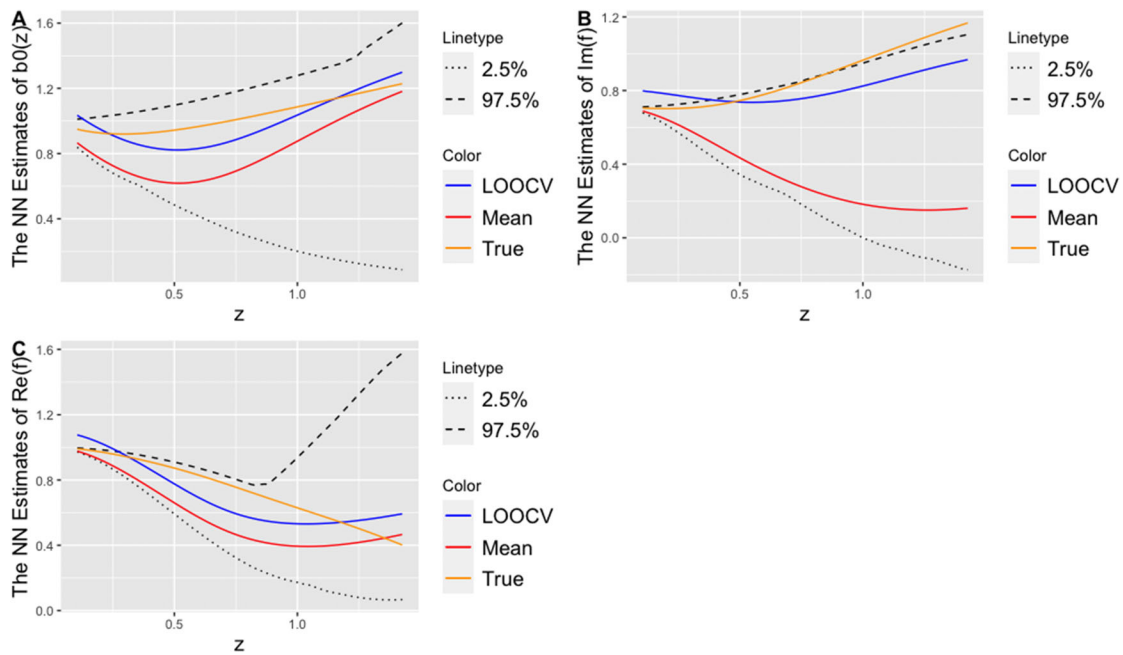
Fig. 9 The posterior mean (red), LOOCV (blue) and the true, using $\omega = 1.003$, (orange) NN-based estimates of the critical collapse functions $b_0(z)$ (**A**), $\text{Im}(f(z))$ (**B**) and $\text{Re}(f(z))$ (**C**) corresponding to the β-solution scenario of the hyperbolic equations of motion. The dotted and dashed lines show, respectively, the lower and upper bounds of the 95% Bayesian credible intervals under Uniform prior
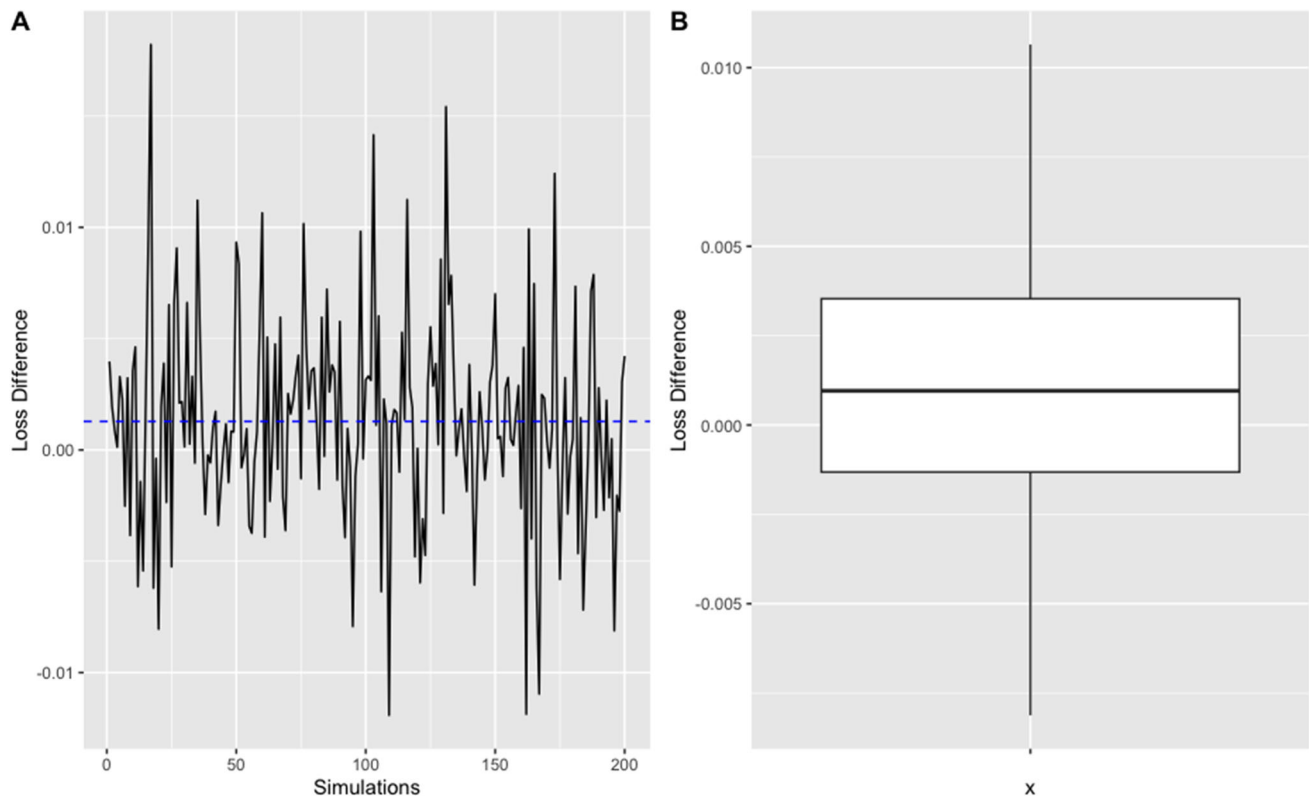
Fig. 10 The trace- and box- plots of the differences between train and test loss values in the NN-based solvers using $L = 200$ SMC samples from the posterior distribution of $\pi(\omega|\mathbf{y})$ when Uniform distribution was used as prior

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**, 337 (1986)
3. D. Christodoulou, Global existence of generalized solutions of the spherically symmetric Einstein scalar equations in the large. Commun. Math. Phys. **106**, 587 (1986)
4. D. Christodoulou, The structure and uniqueness of generalized solutions of the spherically symmetric Einstein scalar equations. Commun. Math. Phys. **109**, 591 (1987)
5. R.S. Hamade, J.H. Horne, J.M. Stewart, Continuous self-similarity and *S*-duality. Class. Quantum Gravity **13**(1996), 2241–2253 (1995). arXiv:gr-qc/9511024
6. R.S. Hamade, J.M. Stewart, The spherically symmetric collapse of a massless scalar field. Class. Quantum Gravity **13**, 497 (1996). arXiv:gr-qc/9506044
7. T. Koike, T. Hara, S. Adachi, Critical behavior in gravitational collapse of radiation fluid: a renormalization group (linear perturbation) analysis. Phys. Rev. Lett. **74**, 5170 (1995). arXiv:gr-qc/9503007
8. L. Alvarez-Gaume, C. Gomez, M.A. Vazquez-Mozo, Scaling phenomena in gravity from QCD. Phys. Lett. B **649**, 478 (2007). arXiv:hep-th/0611312
9. M. Birukou, V. Husain, G. Kunstatter, E. Vaz, M. Olivier, Scalar field collapse in any dimension. Phys. Rev. D **65**, 104036 (2002). arXiv:gr-qc/0201026
10. V. Husain, G. Kunstatter, B. Preston, M. Birukou, Anti-de Sitter gravitational collapse. Class. Quantum Gravity **20**, L23 (2003). arXiv:gr-qc/0210011
11. E. Sorkin, Y. Oren, On Choptuik's scaling in higher dimensions. Phys. Rev. D **71**, 124005 (2005). arXiv:hep-th/0502034
12. J. Bland, B. Preston, M. Becker, G. Kunstatter, V. Husain, Dimension-dependence of the critical exponent in spherically symmetric gravitational collapse. Class. Quantum Gravity **22**, 5355 (2005). arXiv:gr-qc/0507088
13. E.W. Hirschmann, D.M. Eardley, Universal scaling and echoing in gravitational collapse of a complex scalar field. Phys. Rev. D **51**, 4198 (1995). arXiv:gr-qc/9412066
14. J.V. Rocha, M. Tomašević, Self-similarity in Einstein–Maxwell-dilaton theories and critical collapse. Phys. Rev. D **98**(10), 104063 (2018). arXiv:1810.04907 [gr-qc]
15. L. Alvarez-Gaume, C. Gomez, A. Sabio Vera, A. Tavanfar, M.A. Vazquez-Mozo, Critical gravitational collapse: towards a holographic understanding of the Regge region. Nucl. Phys. B **806**, 327 (2009). arXiv:0804.1464 [hep-th]
16. C.R. Evans, J.S. Coleman, Observation of critical phenomena and self-similarity in the gravitational collapse of radiation fluid. Phys. Rev. Lett. **72**, 1782 (1994). arXiv:gr-qc/9402041
17. D. Maison, Non-universality of critical behaviour in spherically symmetric gravitational collapse. Phys. Lett. B **366**, 82 (1996). arXiv:gr-qc/9504008
18. A. Strominger, L. Thorlacius, Universality and scaling at the onset of quantum black hole formation. Phys. Rev. Lett. **72**, 1584 (1994). arXiv:hep-th/9312017
19. E.W. Hirschmann, D.M. Eardley, Critical exponents and stability at the black hole threshold for a complex scalar field. Phys. Rev. D **52**, 5850 (1995). arXiv:gr-qc/9506078
20. A.M. Abrahams, C.R. Evans, Critical behavior and scaling in vacuum axisymmetric gravitational collapse. Phys. Rev. Lett. **70**, 2980 (1993)

21. L. Alvarez-Gaume, C. Gomez, A. Sabio Vera, A. Tavanfar, 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]

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

23. R. Antonelli, E. Hatefi, On critical exponents for self-similar collapse. JHEP **03**, 180 (2020). arXiv:1912.06103 [hep-th]

24. E. Hatefi, A. Kuntz, On perturbation theory and critical exponents for self-similar systems. Eur. Phys. J. C **81**(1), 15 (2021). arXiv:2010.11603 [hep-th]

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

26. E. Hatefi, E. Vanzan, On higher dimensional self-similar axion–dilaton solutions. Eur. Phys. J. C **80**(10), 952 (2020). [arXiv:2005.11646 [hep-th]]

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

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

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

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

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

32. L. Álvarez-Gaumé, C. Gómez, M.A. Vázquez-Mozo, Scaling phenomena in gravity from QCD. Phys. Lett. **B649**, 478–482 (2007). arXiv: hep-th/0611312

33. E. Hatefi, A. Nurmagambetov, I. Park, ADM reduction of IIB on $\mathcal{H}^{p,q}$ to dS braneworld. JHEP **04**, 170 (2013). arXiv:1210.3825

34. E. Hatefi, A. Nurmagambetov, I. Park, N^3 entropy of $M5$ branes from dielectric effect. Nucl. Phys. B **866**, 58–71 (2013). arXiv:1204.2711

35. S. de Alwis, R. Gupta, E. Hatefi, F. Quevedo, Stability, tunneling and flux changing de Sitter transitions in the large volume string scenario. JHEP **11**, 179 (2013). arXiv:1308.1222

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

37. R.S. Hamade, J.H. Horne, J.M. Stewart, Continuous selfsimilarity and *S*-duality. Class. Quantum Gravity **13**, 2241 (1996). arXiv:gr-qc/9511024

38. J. Bland, B. Preston, M. Becker, G. Kunstatter, V. Husain, Dimension dependence of the critical exponent in spherically symmetric gravitational collapse. Class. Quantum Gravity **22**, 5355–5364 (2005)

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

40. L. Álvarez-Gaumé, E. Hatefi, Critical collapse in the axion-dilaton system in diverse dimensions. Class. Quantum Gravity **29**, 025006 (2012). arXiv:1108.0078 [gr-qc]

41. L. Álvarez-Gaumé, E. Hatefi, More on critical collapse of axion-dilaton system in dimension four. JCAP **1310**, 037 (2013). arXiv:1307.1378 [gr-qc]

42. E. Hatefi, A. Hatefi, Estimation of critical collapse solutions to black holes with nonlinear statistical models. Mathematics **10**(23), 4537 (2022). arXiv:2110.07153 [gr-qc]

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

44. E. Hatefi, A. Hatefi, R.J. López-Sastre, Analysis of black hole solutions in parabolic class using neural networks. Eur. Phys. J. C **83**, 623 (2023). arXiv:2302.04619 [gr-qc]

45. E. Hatefi, A. Hatefi, R.J. López-Sastre, Modeling the complexity of elliptic black hole solution in 4D using Hamiltonian Monte Carlo with stacked neural networks. arXiv:2307.14515 [gr-qc]

46. Original reviews are A. Sen, Strong-weak coupling duality in four-dimensional string theory. Int. J. Mod. Phys. A **9**, 3707 (1994). arXiv:hep-th/9402002

47. J.H. Schwarz, Evidence for non perturbative string symmetries. Lett. Math. Phys. **34**, 309 (1995). [arXiv:hep-th/9411178]

48. M.B. Green, J.H. Schwarz, E. Witten, *Superstring Theory*, vols. I, II (Cambridge University Press, Cambridge, 1987)

49. J. Polchinski, *String Theory*, vols. I, II (Cambridge University Press, Cambridge, 1998)

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

51. R.Y. Rubinstein, D.P. Kroese, *Simulation and the Monte Carlo Method*, vol. 707 (Wiley, New York, 2011)

52. C.P. Robert, G. Casella, *Monte Carlo Statistical Methods*, vol. 2 (Springer, New York, 1999)

53. P. Del Moral, A. Doucet, A. Jasra, Sequential Monte Carlo samplers. J. R. Stat. Soc. Ser. B Stat. Methodol. **68**(3), 411–436 (2006)

54. J.M. Bernardo, M.J. Bayarri, J.O. Berger, A.P. Dawid, D. Heckerman, A.F.M. Smith, M. West, P. Del Moral, A. Doucet, A. Jasra, Sequential Monte Carlo for Bayesian computation. Bayesian Stat. **8**, 1–34 (2011)

55. C.M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006)

56. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, Cambridge, 2016)

57. I.E. Lagaris, A. Likas, D.I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations. IEEE Trans. Neural Netw. **9**(5), 987–1000 (1998)

58. J. Salvatier, T.V. Wiecki, C. Fonnesbeck, Probabilistic programming in Python using PyMC3. PeerJ Comput. Sci. **2**, e55 (2016)

59. F. Chen, D. Sondak, P. Protopapas, P. Mattheakis, M. Liu, S. Agarwal, D. Di Giovanni, NeuroDiffEq: a Python package for solving differential equations with neural networks. J. Open Source Softw. **5**, 1931 (2020)