AdS/PINN - bulk theory from boundary data: AdS/DL correspondence with flexible functional forms for bulk action

Minseong Kim*

(Dated: July 15, 2025 (v2))

Abstract

The AdS/DL (AdS/deep learning) correspondence was implemented as the AdS/DBM correspondence in Hashimoto (2019), the key idea being that DBM (deep Boltzmann machine) states and energy admit physical interpretations such that as long as we provide the right boundary data, bulk physics can be recovered by training DBM. However, such interpretations are possible only for a very restrictive class of theories such as non-interacting theories, and even without considering feasibility of physical interpretations, DBM is no longer used in practice due to its poor training performances. This paper extends the idea of recovering bulk physics from boundary data more generally to any boundary theory within the supervised physics-informed neural network framework - AdS/PINN. Expensive self-supervised (unsupervised) training measures such as KL divergence are avoided by the score mechanism that leverages the known boundary theory itself (supervised PINN). AdS/PINN opens up for directly studying the holographic bulk duals of weakly coupled theories as well as holography beyond AdS/CFT.

* mkimacad@gmail.com; ORCiD:0000-0003-2115-081X

CONTENTS

I. INTRODUCTION

In recent years, AdS/CFT has been analyzed within the AdS/DL correspondence [1], with supervised and self-supervised generative versions [2]. Conventional supervised learning cases of AdS/DL tend to be fairly limited - one does not attempt to completely recover the bulk theory and the bulk metric from the boundary theories by deep learning and instead attempts to recover parts of the bulk metric and a few bulk observables that remain fairly constrained.

Although the AdS/DBM correspondence in [2] removes these constraints by understanding the DBM (deep Boltzmann machine) weights as corresponding to the bulk metric, the DBM nevertheless limits the possible functional forms of the bulk action for the purpose of physical interpretations. As a theoretical demonstration that neural networks can provide entire holographic physics from boundary data, AdS/DBM has severe limits.

Furthermore, DBM has been considered to be largely deprecated and is no longer actively used. Thus, even when DBM is not used as a physics-interpretable network such that it allows for any functional form of bulk action, it is severely limited.

This paper extends and makes fully operational the spirit of obtaining holographic bulk physics from boundary data in AdS/DBM - AdS/PINN, where PINN stands for physics-informed neural networks. The main goal is to show a proof-of-concept (PoC) or a framework demonstrating that as long as any boundary theory data are available, neural networks can be trained to provide interpretable holographic theories in the form of bulk action $S(\phi)$, where ϕ is bulk field.

Training-wise, AdS/PINN improves upon AdS/DBM by moving away from expensive unsupervised learning techniques (training DBM) to supervised PINN learning techniques (score mechanism) based on the physics-informed neural networks framework (where target scores are informed by known boundary theory physics) while maintaining the spirit of obtaining the bulk from boundary theory data alone.

Physics-wise justifications for exploring AdS/PINN mainly come from the disconnections between theories explored in the UV scale (many of them being conformal field theories), those explored in the IR scale (usually quantum field theories that break scale invariance) and those explored at the classical level. Even the question of how classicality emerges is yet to be completely understood. Holography offers a natural strategy for bridging these theories due to holographic renormalization (see Section II C), but the known holography, AdS/CFT, is constrained to conformal boundary field theories. If we take the bulk theory to be our observable universe, then the current data suggest that the universe is not close to AdS space. AdS/PINN, allowing for perturbations of CFTs in boundary data, is a powerful tool for studying these cases, especially in light of braneworld scenarios where the observable universe is mostly constrained to some effectively four-dimensional boundary without much observable access to existing additional dimensions [3, 4].

There are other physics-wise justifications of AdS/PINN that are shared with existing supervised learning cases of AdS/DL (which are explored in Section II A), such as finding unknown metric components of the largely known asymptotically AdS space, and in these instances AdS/PINN offers a more flexible approach.

It is also noted that AdS/PINN is not limited to the proof of concept in this paper. Relativistic constraints can also be imposed and even general-relativistic ones with explicit

FIG. 1: The goal is to find out the physical law behind the black box, with inputs given by (x_i, v_i) and outputs given by (x_f, v_f), with x understood as position and v understood as velocity.

background metric learning, though training such neural networks would inevitably be complex, requiring significant computational resources.

We briefly note the quasi-fundamental limitation of AdS/PINN - the question of how numerically learned action may be converted to something like Lagrangian cannot yet be answered - while we know what happens in the bulk black box, a tractable human-readable analysis requires further processing of learned bulk action data.

In Section II, we revisit the conventional uses of deep learning in holographic gravity or AdS/DL. The implicit holographic renormalization rationale behind Section III is justified in Section II C. The details of AdS/PINN are then presented in Section III, with the proof-of-concept (PoC) code provided as stated in the appendix.

II. CONVENTIONAL ADS/DL REVISITED

A. Supervised AdS/DL

In [5], a simple classical example is used to outline the principles behind AdS/DL, which we partially replicate. For the setup described in Figure 1, conventional supervised learning with inputs (x_i, v_i) and outputs (x_f, v_f) can be done such that for some (x_i, v_i) we get predictions (x_f, v_f). However, physical interpretability of this neural network black box is a different matter. Now the following discretized flexible equation of motion is assumed:

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + v^{(k)}\,\Delta t\,, \\ v^{(k+1)} &= v^{(k)} + f(x^{(k)}, v^{(k)})\,\Delta t\,, \end{aligned} \tag{1}$$

Input (x_i, v_i) corresponds to $\mathbf{x}^{(0)} = (x^{(0)}, v^{(0)})^T$, while output (x_f, v_f) corresponds to $\bar{\mathbf{x}}^{(\text{out})} = (x^{(\text{end})}, v^{(\text{end})})^T$ ideally. This system of equations is cast into a form resembling a neural network:

$$\mathbf{x}^{(k+1)} = \varphi^{(k)} \left(W^{(k)} \mathbf{x}^{(k)} \right) , \tag{2}$$

where

$$W^{(k)} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} , \qquad \varphi^{(k)} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b + f(x^{(k)}, v^{(k)}) \, \Delta t \end{pmatrix} . \tag{3}$$

and training is now modified to learning activation function φ, which is also about learning $f(x^{(k)}, v^{(k)})$ at each layer k. We can safely assume that f is modeled as a neural network, with the loss function is changed to:

$$L = \frac{1}{n_{\text{batch}}} \sum_{\text{batch}} |\bar{\mathbf{x}}^{(\text{out})} - \mathbf{x}^{(T)}| + L_{\text{reg}} \, , \tag{4}$$

Typically, it is not an activation function that is learned - the setup of linear trainable weights with fixed activation functions is more usual, as is the case for the AdS/DL correspondence in [1]. There, it is metric that is a black box, with metric determining field equations of motion completely, assuming the large-N limit. A flexible functional form for (asymptotically) AdS metric goes as:

$$ds^2 = -f(\eta) dt^2 + d\eta^2 + g(\eta)(dx_1^2 + \cdots + dx_{d-1}^2) \tag{5}$$

The assumed scalar field theory goes as:

$$S = \int d^{d+1}x \sqrt{-\det g} \left[-\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2 - V(\phi) \right] . \tag{6}$$

Discretizing the resulting equations of motion, we obtain:

$$\begin{aligned} \phi(\eta + \Delta\eta) &= \phi(\eta) + \Delta\eta \, \pi(\eta) \\ \pi(\eta + \Delta\eta) &= \pi(\eta) - \Delta\eta \left(h(\eta)\pi(\eta) - m^2 \phi(\eta) - \frac{\delta V(\phi)}{\delta \phi(\eta)} \right) \end{aligned} \tag{7}$$

where:

$$\pi \equiv \partial_\eta \phi, \quad h(\eta) \equiv \partial_\eta \log \sqrt{f(\eta) g(\eta)^{d-1}} \tag{8}$$

Therefore, the goal is to train weight $h(\eta)$ at each discrete η, given input $(\phi(\eta_{ini}), \pi(\eta_{ini}))$ at $\eta_{ini} \approx \infty$ matched with binary output $y \in \{0, 1\}$ that labels whether the boundary condition at $\eta_{fin} \approx 0$ is satisfied, output zero labeling satisfaction. The boundary condition satisfaction refers to $F = 0$, with F being used as the final output activation function:

$$F(\phi, \pi) \equiv \left[\frac{2}{\eta} \pi - m^2 \phi - \frac{\delta V(\phi)}{\delta \phi} \right]_{\eta = \eta_{\text{fin}}} \tag{9}$$

AdS/CFT	AdS/DBM
Bulk radial coordinate z	Hidden layer label k
QFT source $J(x)$	Input value v_i
Bulk field $\phi(x,z)$	Hidden variables $h_i^{(k)}$
QFT generating functional $Z[J]$	Probability distribution $P(\mathbf{v})$
Bulk action $S[\phi]$	Energy function $\mathcal{E}(\mathbf{v},\mathbf{h}^{(\mathbf{0<k\leq f})})$

TABLE I: From [2].

As a neural network form with $\mathbf{x}^{(\mathbf{k})} = (\phi^{(k)}, \pi^{(k)})^T$ and $\eta^{(k)} = \eta_{ini} + k\Delta\eta$,

$$\begin{aligned} W^{(n)} &= \begin{pmatrix} 1 & \Delta\eta \\ \Delta\eta\, m^2 & 1 - \Delta\eta\, h(\eta^{(n)}) \end{pmatrix}, \\ \varphi_1(x_1) &= x_1 \\ \varphi_2(x_2) &= x_2 + \Delta\eta\, \tfrac{\delta V(x_1)}{\delta x_1}, \\ y(\mathbf{x}^{(\mathbf{0})}) &= F(\varphi(W^{(N-1)}\varphi(W^{(N-2)}\cdots\varphi(W^{(0)}\mathbf{x}^{(\mathbf{0})})))). \end{aligned} \tag{10}$$

A loss function example goes as:

$$L = \frac{1}{n_{batch}} \sum_{batch} |\bar{y} - y| + L_{reg} \tag{11}$$

where $\bar{y}$ is actual output data. $h(\eta)$ is then updated via backpropagation and some variant of gradient descent. We note several papers that are in a similar vein [6–9].

B. Hashimoto (2019): (generative) AdS/DL as deep Boltzmann machines (DBM)

So far, the discussion has been about filling the missing details of the bulk theory. However, ideally, deep learning should try to learn the entire bulk theory from boundary theory data if holography is true. This is what the AdS/DBM of Hashimoto (2019) [2] attempted to demonstrate - that we can translate neural networks into interpretable bulk physics. The cost involved for DBM is that we now move from conventional supervised learning involving input-output pairs to training generative models. What is learned in AdS/DBM are boundary theory partition functions with additional gravity-inspired regularization, and this is why self-supervised generative models are involved.

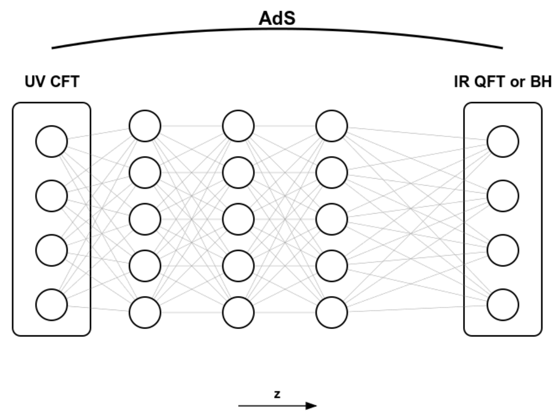

FIG. 2: Generative AdS/DL as deep Boltzmann machine (DBM). AdS space here is asymptotically AdS space. It resembles a fully connected feedforward network, but each node (neuron) output $h_i^{(k)}$ (node $0 \leq i < w$ at layer $0 \leq k \leq \ell$) is not deterministically dependent on nodes of preceding layers. Rather, DBM weights determine the probability by which some neuron value arises. In this paper, the IR boundary can be perturbed to break scale invariance, modeling some effective field theory (EFT). AdS/PINN can map holographic responses to such perturbations.

DBMs are universal approximators and therefore can learn any partition function. However, if we are to translate weights of DBM into interpretable physics, the map provided by [2] is restricted to the only limited class of theories, mostly around non-interacting theories. **This restriction is what we intend to eliminate in AdS/PINN.** After all, why not make neural networks learn and output the physics in form of action? There is no reason to restrict translations to neural network weights.

We now briefly describe AdS/DBM. The DBM energy $\mathcal{E}(\mathbf{v})$ is given by, with v_i being part of visible layer $\mathbf{v}$:

$$\mathcal{E}(\mathbf{v}) = \sum_{i,j} w_{ij}^{(0)} v_i h_j^{(1)} + \sum_{k=1}^{N-1} \left[\sum_{i,j} w_{ij}^{(k)} h_i^{(k)} h_j^{(k+1)} \right] \tag{12}$$

with probability of $\mathbf{v}$ then given as:

$$P(\mathbf{v}) \propto \sum_{h_i^{(k)}} \exp\left(-\mathcal{E}(\mathbf{v})\right) \tag{13}$$

It is on this probability that AdS/DBM trains the bulk on, though sampling boundary data as to match underlying probability distributions (which are not given and have to be determined from the partition functions) as well as fitting probability distributions in the bulk is significantly more difficult relative to supervised learning.

C. AdS/DL as holographic renormalization

It is natural to interpret AdS/DL in the context of holographic renormalization, where the UV CFT fixed point is reached with the inverse renormalization flow of $z \to 0$, and the IR CFT fixed point is reached with the renormalization flow of $z \to \infty$. Alternatively, we may introduce $z = z_\ell$ in place of $z \to \infty$, with which we understand Euclidean partition function as follows [10, 11] (note that [11] is mostly followed without content-wise modifications):

$$
\begin{aligned}
Z[J] &= \int \mathcal{D}\tilde{\phi} \int \mathcal{D}\phi|_{z>z_\ell} e^{-S|_{z>z_\ell}} \int \mathcal{D}\phi|_{z<z_\ell}\, e^{-S|_{z<z_\ell}} \\
&\equiv \int \mathcal{D}\tilde{\phi}\, \Psi_{IR}[\tilde{\phi}; z_\ell] \Psi_{UV}[J, \tilde{\phi}; \epsilon, z_\ell].
\end{aligned} \tag{14}
$$

where $\phi(x,\epsilon) = \epsilon^{d-\Delta} J(x)$, $d+1$ being AdS spacetime dimension, $\Delta = \frac{d}{2} + \sqrt{\frac{d^2}{4} + m^2}$ and boundary conditions $\phi(x, z_\ell) = \tilde{\phi}(x)$. We can further decompose Ψ_{IR} and Ψ_{UV} as:

$$\Psi_{UV}[J, \tilde{\phi}; \epsilon, z_\ell] = \int \mathcal{D}\phi_\epsilon\, \Psi_{UV}[J, \phi_\epsilon, \tilde{\phi}; \epsilon, z_\ell] \tag{15}$$

$$\Psi_{UV}[J, \phi_\epsilon, \tilde{\phi}; \epsilon, z_\ell] = \int_{\phi(x,\epsilon)=\phi_\epsilon(x)}^{\phi(x,z_\ell)=\tilde{\phi}(x)} \mathcal{D}\phi\, e^{-S|_{z<z_\ell}} \tag{16}$$

$$\Psi_{IR}[\tilde{\phi}; z_\ell] = \int \mathcal{D}\phi_f\, \Psi_{IR}[\tilde{\phi}, \phi_f; z_\ell] \tag{17}$$

$$\Psi_{IR}[\tilde{\phi}, \phi_f; z_\ell] = \int_{\phi(x,z_\ell)=\tilde{\phi}(x)}^{\phi(x,z_f)=\phi_f} \mathcal{D}\phi\, e^{-S|_{z>z_\ell}} \tag{18}$$

$$\tilde{P}[J, \phi_\epsilon, \tilde{\phi}, \phi_f; \epsilon, z_\ell] = \Psi_{UV}[J, \phi_\epsilon, \tilde{\phi}; \epsilon, z_\ell] \Psi_{IR}[\tilde{\phi}, \phi_f; z_\ell] \tag{19}$$

$$Z[J] = \int \mathcal{D}\phi_\epsilon \mathcal{D}\tilde{\phi} \mathcal{D}\phi_f\, \tilde{P}[J, \phi_\epsilon, \tilde{\phi}, \phi_f; \epsilon, z_\ell] \tag{20}$$

In [11], this renormalization framework is pushed further in the $z_\ell \to 0$ limit to prove the equivalence of two AdS/CFT dictionaries, GKPW [12, 13] and BDHM dictionaries [14]. For the purpose of this paper, this is unnecessary.

III. PROOF-OF-CONCEPT FOR ADS/PINN

A. General mechanism

Training-wise, expensive KL divergence losses that match partition functions are avoided and techniques from physics-informed neural networks are adopted instead. For brevity, neural network names used in the code are abbreviated:

- **PathSolverCNN** is abbreviated to **PathCNN**,
- **StabilizedClassicalPathSolverNN** can be abbreviated to **ClassicalPathNN** and is further abbreviated to **ClassicalNN**,
- **ActionPerturbationCNN** is abbreviated to **ActionCNN**,
- **CFTSampleGenerator** is abbreviated to **Sampler**,
- and CNN refers to convolution neural networks, NN refers to conventional residual feedforward networks.

The problem is broken down into multiple neural networks:

1. **Classical path resolver**: The problem of finding the classical bulk field that matches boundary field data, which is used for calculating bulk action. A dedicated convolutional neural network (CNN), **PathCNN**, encodes UV and IR boundary data and feeds the encoded data to **ClassicalNN**, which, along with the input of radial coordinate z, produces the classical bulk field that matches the given boundary field data.

2. **Action resolver** The problem of calculating the action for each bulk field, which is handled by **ActionCNN**, a CNN that takes in boundary field data and produces action perturbations to the reference theory.

3. **Score mechanism (loss)**: The neural networks are connected by the loss function, with the main loss term being the score mechanism:

$$L_{\text{score}} = \frac{1}{N}\sum_{i=1}^{N}\left(||\nabla_{\phi_{\text{UV}}} S_{\text{model}} - \nabla_{\phi_{\text{UV}}} S_{\text{CFT}}||^2 + ||\nabla_{\phi_{\text{IR}}} S_{\text{model}} - \nabla_{\phi_{\text{IR}}} S_{\text{target}}||^2\right) \tag{21}$$

where S_{model} would refer to the action provided by ActionPerturbationCNN summed with the action of the reference theory.

4. **Boundary field data sampler**: The score mechanism requires samplers for UV and IR boundary fields, which are provided by **Sampler**, initially generating a random white noise and filtering out according to the two-point correlation function of 2D massless free scalar field theory scaling as $1/k^2$ in momentum space (which does not take into account divergences, but this is irrelevant here). The same sampler is used for generating IR boundary data for simplicity - and this is not a problem in practice, especially given that samplers do not have to be faithful to partition functions. The score mechanism learns partition functions, not samplers.

5. **Theoretical boundary actions**: these are provided by **free_scalar_action** (UV boundary) and **ir_act** (IR boundary). They provide S_{CFT} (UV) and S_{target} (IR) in the score mechanism.

6. **Perturbation on the IR boundary**: this is controlled by parameter p. Parameter p is also an input to **ActionCNN** and **ClassicalCNN**. Note that the code can be modified to incorporate other perturbations and p can be a vector - see Section III B.

7. The UV boundary theory chosen in the PoC code is the two-dimensional free massless scalar field theory, a trivial CFT. Now to slightly less fundamental details.

8. **Reference theory**: The reference base theory that the code (in particular, **ActionCNN**) uses is the three-dimensional flat-space free massless scalar field theory, which is not the actual holographic bulk theory. This is an intentional choice to avoid the impression that AdS/PINN simply learns from the existing regularization choice. This is why we should expect significant action perturbation values even for $p = 0$, because action perturbations do not refer to the perturbations (**act_pert** or **action_pert**) from the holographic bulk theory at $p = 0$. The reference theory only exists as an initial guess to temporarily guide neural networks and to avoid large calculations that create numerical and training issues.

9. **Reference path**: its purpose similar to the reference base theory - in the PoC code, this reference path is a simple linear interpolation of boundary fields. Path deviations

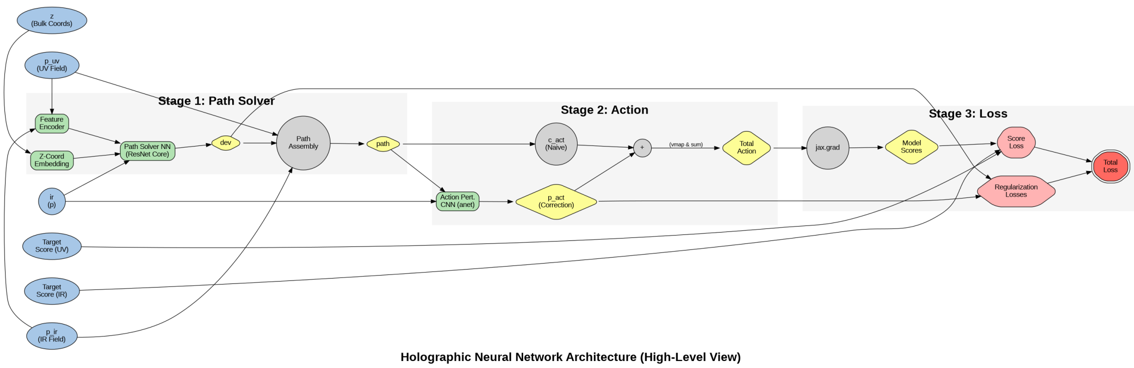

FIG. 3: The high-level diagram of the AdS/PINN PoC network structure - the horizontal view. See Figure 4 in case of the readability issue.

dev then are about how the actual bulk field deviates from this reference bulk field.

For the high-level diagram of the AdS/PINN PoC, see Figure 3, which visualizes the details above.

B. Perturbation examples

In the code, the IR boundary is perturbed from CFT such that it gains mass with perturbation parameter p, which relates to mass m^2, since the IR boundary action is perturbed as:

$$S_{\mathrm{IR}}[\phi, p] = \int d^2x \, \frac{1}{2}(\nabla\phi)^2 + \frac{p}{2}\int d^2x \, \phi(x)^2 \tag{22}$$

Therefore, the code studies how the introduction of mass in the IR theory must perturb the initial holographic bulk theory - 'action perturbation' - in order to satisfy action data in two boundaries.

Perturbations are not restricted to breaking scale invariance of boundaries - even within AdS/CFT, as we noted in Section II, there are cases where the full bulk gravity solution is unknown, and they can be studied as perturbations over the well-known examples. There is nothing stopping us from generalizing the PoC code, and p can also be a vector.

C. Score mechanism and action

The score mechanism has a slight deviation from the conventional fixed background AdS/CFT, because the classical on-shell condition is imposed for every boundary field data

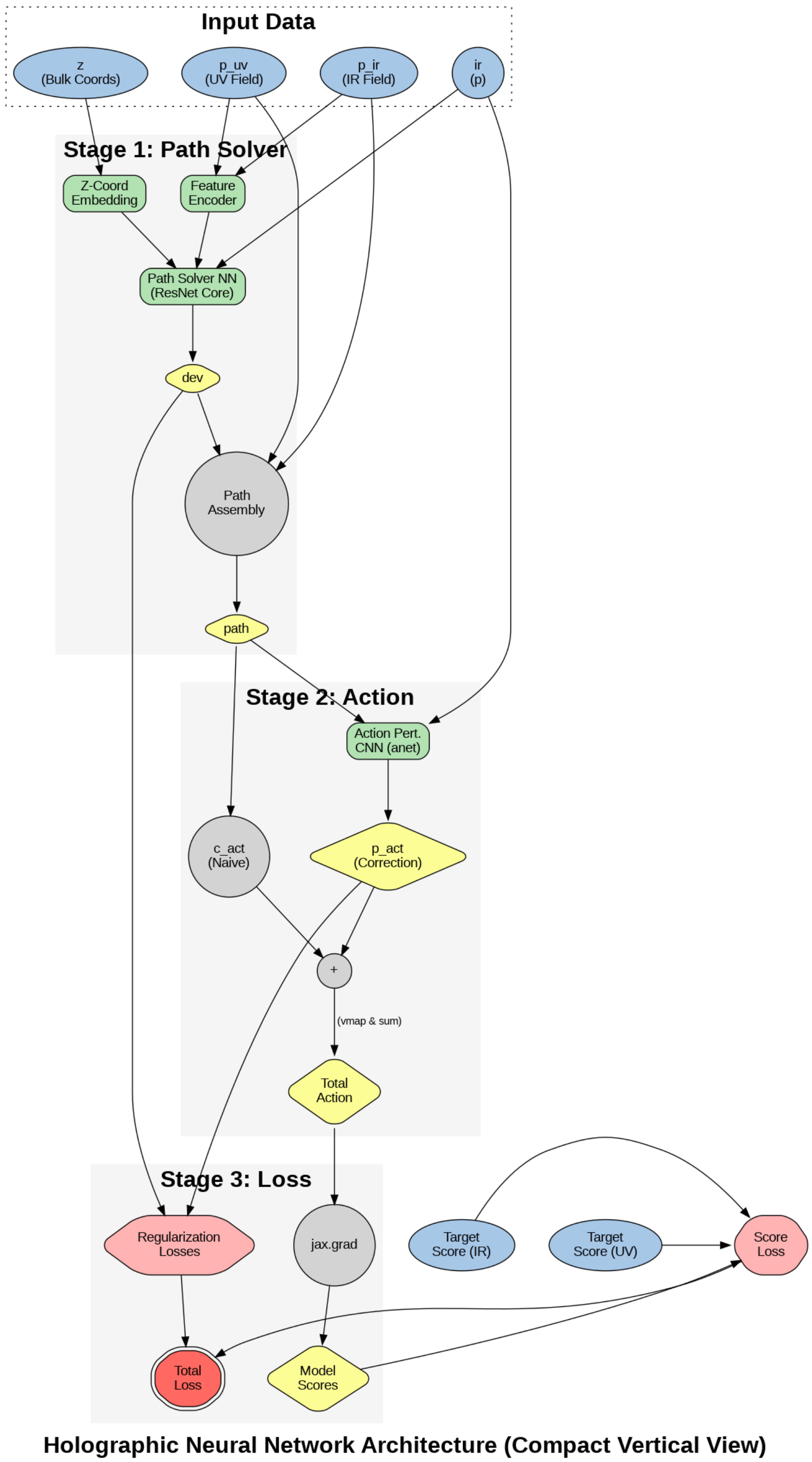

FIG. 4: The high-level diagram of the AdS/PINN PoC network structure - the vertical view.

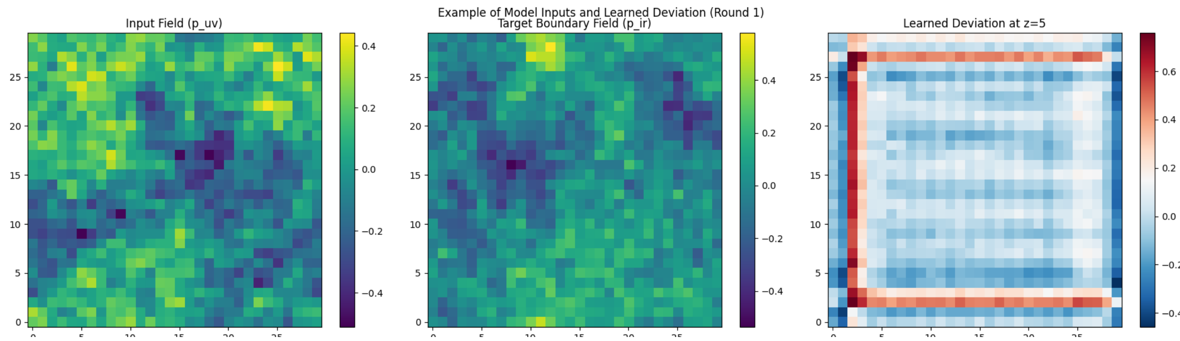

FIG. 5: An example trained and sampled field data result at perturbation parameter $p = 5.0$ from the PoC code.

sampled:

$$\nabla_{\phi_{\mathrm{b1}}} S_{\mathrm{bulk}}|_{\phi_{b2}} \approx \nabla_{\phi_{\mathrm{b1}}} S_{\mathrm{b1}} \tag{23}$$

where $b1, b2$ refer to different boundaries. In the conventional AdS/CFT, boundary conditions (Dirichlet or Neumann boundary conditions) along with some boundary data on the UV boundary determine the entire classical on-shell bulk field, with some variations available. Therefore, the question raised by Equation (23) is trivial in the conventional fixed-background AdS/CFT.

This does not mean that boundary fields that are classically off-shell never arise, since the off-shell condition here does not have to do with relativity. After all, even the bulk theory must be some quantum theory if we probe deep enough, which induces different classical limits depending on states. The code and the analysis of this paper therefore considers how classical gravity and the bulk must react in case it sees any sampled boundary field data. This is a virtue, not weakness, especially in light of the strong/weak duality and weakly-coupled boundary theories (which have strongly-coupled holographic bulk) in holographic gravity.

D. On expected neural network generalization error

In the past decades, a common theoretical criticism of over-parameterized neural networks was that they are bound to suffer from generalization error eventually. The recent decade saw the reversal of this viewpoint with the rise of neural tangent kernel (NTK) analysis [15, 16] and the F-principle [17]. The idea has been that infinite-width neural networks with gradient descent can be analyzed as linear models, with the networks exponentially converging to interpolants of minimum RKHS norm, which roughly translate to the ones with least complexity consistent with training data. The F-principle further states that neural networks learn low frequencies much faster than high frequencies. This allows the expected generalization error to become smaller as the training data accumulate. In essence, this implies that practical applicability of AdS/DL is tied to complexity of physics involved, as over-parameterized neural networks are biased toward less complexity.

There are some remaining issues: neural networks nevertheless suffer whenever they generalize beyond the interpolation domain [18]. There have been attempts to fix this issue, but the answer is yet to come out satisfactorily.

IV. CONCLUSION

AdS/PINN generalizes AdS/DBM (AdS/DL in terms of DBM) of [2] by allowing for unrestricted functional forms of the bulk theory action for the purpose of physically interpreting neural networks. AdS/PINN learns how classical gravity and the bulk react to different boundary field data by the score mechanism of Equation (21). The spirit of AdS/DBM that the bulk is learned from boundary field data is extended without invoking expensive unsupervised learning and divergence measures such as KL divergence that tries to match target probability distributions.

Boundary theories are not particularly constrained in AdS/PINN, and this allows us to study how far holography can go as perturbations of AdS/CFT. Even within AdS/CFT, this allows us a probe into the holographic dual of a weakly-coupled boundary theory as a perturbation over some strongly-coupled boundary theory for which the bulk solution is already known. The score mechanism of Equation (21) is designed to handle cases where gravity strongly fluctuates and the classical background varies depending on excited boundary field states.

There are several constraints that are left out in the proof-of-concept in this paper, such as relativity constraints for the bulk and how gauge constraints are to be handled, which do not appear in scalar theories. Their implementations are not theoretically prohibited as they are simply regularization or loss terms. For now, we defer these questions to future works.

[1] Koji Hashimoto, Sotaro Sugishita, Akinori Tanaka, and Akio Tomiya, "Deep learning and the AdS/CFT correspondence," Physical Review D **98** (2018), 10.1103/physrevd.98.046019.

[2] Koji Hashimoto, "AdS/CFT correspondence as a deep Boltzmann machine," Physical Review D **99** (2019), 10.1103/physrevd.99.106017.

[3] Lisa Randall and Raman Sundrum, "Large mass hierarchy from a small extra dimension," Physical Review Letters **83**, 3370–3373 (1999).

[4] Lisa Randall and Raman Sundrum, "An alternative to compactification," Physical Review Letters **83**, 4690–4693 (1999).

[5] Mugeon Song, Maverick S. H. Oh, Yongjun Ahn, and Keun-Young Kima, "AdS/Deep-Learning made easy: simple examples," Chinese Physics C **45**, 073111 (2021).

[6] Tetsuya Akutagawa, Koji Hashimoto, and Takayuki Sumimoto, "Deep learning and AdS/QCD," Physical Review D **102** (2020), 10.1103/physrevd.102.026020.

[7] Yu-Kun Yan, Shao-Feng Wu, Xian-Hui Ge, and Yu Tian, "Deep learning black hole metrics from shear viscosity," Phys. Rev. D **102**, 101902 (2020).

[8] Kai Li, Yi Ling, Peng Liu, and Meng-He Wu, "Learning the black hole metric from holographic conductivity," Phys. Rev. D **107**, 066021 (2023).

[9] Byoungjoon Ahn, Hyun-Sik Jeong, Keun-Young Kim, and Kwan Yun, "Holographic reconstruction of black hole spacetime: machine learning and entanglement entropy," Journal of High Energy Physics **2025** (2025), 10.1007/jhep01(2025)025.

[10] Idse Heemskerk and Joseph Polchinski, "Holographic and Wilsonian renormalization groups," Journal of High Energy Physics **2011** (2011), 10.1007/jhep06(2011)031.

[11] Daniel Harlow and Douglas Stanford, "Operator Dictionaries and Wave Functions in AdS/CFT and dS/CFT," (2011), arXiv:1104.2621 [hep-th].

[12] Edward Witten, "Anti-de Sitter space and holography," Adv. Theor. Math. Phys. **2**, 253–291 (1998).

[13] S.S. Gubser, I.R. Klebanov, and A.M. Polyakov, "Gauge theory correlators from non-critical string theory," Physics Letters B **428**, 105–114 (1998).

[14] Tom Banks, Michael R. Douglas, Gary T. Horowitz, and Emil Martinec, "AdS Dynamics

from Conformal Field Theory,” (1998), arXiv:hep-th/9808016 [hep-th].

[15] Arthur Jacot, Franck Gabriel, and Clément Hongler, “Neural tangent kernel: convergence and generalization in neural networks,” in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS’18 (Curran Associates Inc., Red Hook, NY, USA, 2018) p. 8580–8589.

[16] Jaehoon Lee, Lechao Xiao, Samuel S Schoenholz, Yasaman Bahri, Roman Novak, Jascha Sohl-Dickstein, and Jeffrey Pennington, “Wide neural networks of any depth evolve as linear models under gradient descent,” Journal of Statistical Mechanics: Theory and Experiment **2020**, 124002 (2020).

[17] Zhi-Qin John Xu Zhi-Qin John Xu, Yaoyu Zhang Yaoyu Zhang, Tao Luo Tao Luo, Yanyang Xiao Yanyang Xiao, and Zheng Ma Zheng Ma, “Frequency principle: Fourier analysis sheds light on deep neural networks,” Communications in Computational Physics **28**, 1746–1767 (2020).

[18] Liu Ziyin, Tilman Hartwig, and Masahito Ueda, “Neural networks fail to learn periodic functions and how to fix it,” in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS ’20 (Curran Associates Inc., Red Hook, NY, USA, 2020).

Appendix A: Code

Available at `https://mkimacad.github.io/ads_vaebm/codes/adsdl.py` as well as an online supplementary material to this paper.