Letter

# Synaptic field theory for neural networks

Donghee Lee,[*] Hye-Sung Lee,[†] and Jaeok Yi[‡]

*Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea*

(Received 25 March 2025; accepted 25 July 2025; published 15 August 2025)

Theoretical understanding of deep learning remains elusive despite its empirical success. In this study, we propose a novel "synaptic field theory" that describes the training dynamics of synaptic weights and biases in the continuum limit. Unlike previous approaches, our framework treats synaptic weights and biases as fields and interprets their indices as spatial coordinates, with the training data acting as external sources. This perspective offers new insights into the fundamental mechanisms of deep learning and suggests a pathway for leveraging well-established field-theoretic techniques to study neural network training.

DOI: 10.1103/bhwt-j2rq

*Introduction.* The remarkable success of deep learning in a wide range of applications has triggered intense efforts to understand its core mechanisms. While deep neural networks are highly effective at extracting rich features, the fundamental reasons for their success remain elusive. Numerous strategies have been proposed to clarify these mysteries; for general reviews, see Refs. [1–4]. One of these strategies is to utilize field theory, which serves as a fundamental framework in physics.

Field-theoretic approaches have been proposed in numerous studies [5–53]. For instance, in Refs. [6–9], the authors propose that deep neural networks reflect the structure of the AdS/CFT correspondence by using neural networks to learn and reconstruct the AdS metric from data. The neural network/Gaussian process correspondence [54–62] has prompted numerous field-theoretic studies, and finite-width effects have been extensively investigated using various approaches [10–28]. There are other intriguing works, too [29–53]. Among these studies, some have emphasized the role of symmetry [28,31–34], while others have applied statistical physics [26–29,35–42]. Especially, the connection to renormalization group transformations has been explored in several works [26,27,35–42].

Krippendorf and Spannowsky (KS) made an important observation, identifying a duality connecting neural network and cosmological dynamics [5]. To substantiate this duality, they took an effective field theory (EFT) approach, and demonstrated that the time evolution of the network's outputs can be mapped onto that of a cosmological system in the limit where the neural tangent kernel (NTK) becomes constant [62].

In this regime, they related cosmological parameters, such as the Hubble parameter, to training parameters of the network, and pointed out a duality between neural networks and de Sitter (dS) space.[1]

Nevertheless, we must develop a more fundamental theory that treats the parameters of neural networks—synaptic weights and biases—directly for various reasons. Here, we introduce a field-theoretic framework that elevates them—the network's fundamental degrees of freedom—to dynamical fields. A theory formulated in terms of these neural network parameters can be viewed as a UV description of the training dynamics.

Because the parameters are numerous, their collective behavior is naturally captured by a continuum description. We therefore construct a "synaptic field theory," the first formalism to treat synapses explicitly as fields. Although taking the continuum limit introduces technical challenges, the conceptual bridge between neural networks and field theory remains clear and robust. This framework brings the full machinery of field theory to neural network analysis and promises new insights into network dynamics.

*Neural network theory and previous works.* A deep neural network (DNN) comprises layers of neurons, each connected by weights $W_{ij}^{(m)}$ and biases $b_i^{(m)} \equiv W_{i0}^{(m)}$. For $(m+1)$th layer, the neuron $h_i^{(m+1)}$ depends on all neurons

[*]Contact author: dhlee641@kaist.ac.kr
[†]Contact author: hyesung.lee@kaist.ac.kr
[‡]Contact author: wodhr1541@kaist.ac.kr

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.*

[1]Although KS used the term "vacuum energy dominated universe," we use dS space for simplicity. While a universe dominated by vacuum energy can indeed be described as dS space, not all dS spacetimes correspond to such a Universe.

$h_j^{(m)}$ in the previous layer according to

$$h_i^{(m+1)} = \sigma\left(\sum_j W_{ij}^{(m)} h_j^{(m)}\right), \tag{1}$$

where $\sigma$ is a nonpolynomial activation function [63]. Here, the sum over $j$ runs from 0 to $N$, and $h_0^{(m)} = 1$ for all $m$ which runs from 0 to $M$. Iterating Eq. (1) from the input (initial layer) $X = h^{(0)}$ to the output (final layer) $Z = h^{(M)}$ defines the forward pass of the network.

Training a DNN involves adjusting $W_{ij}^{(m)}$ to minimize the quadratic cost function

$$C = \sum_{i,l}\left(Y_i^{[l]} - Z_i^{[l]}\right)^2, \tag{2}$$

where $(X_i^{[l]}, Y_i^{[l]})$ are the training inputs and desired outputs, and $Z_i^{[l]}$ is the DNN prediction for $X_i^{[l]}$. The index $l$ labels individual training examples. Gradient descent updates weights and biases according to

$$\Delta W_{ij,T}^{(m)} = -\eta \frac{\partial C}{\partial W_{ij}^{(m)}}, \tag{3}$$

where $\Delta W_T = W_{T+1} - W_T$, with $W_T$ representing the value of $W$ at training step $T$, and $\eta$ denoting the step size. The derivatives on the right-hand side of the equation are evaluated at the weights and biases from the previous training step $T$. Interpreting the discrete step $T$ as a continuous time $t$ yields the differential equation

$$\dot{W} = -\eta \frac{\partial C}{\partial W}, \tag{4}$$

with $W$ representing all $W_{ij}$.

Although sometimes referred to as an equation of motion for gradient descent, Eq. (4), cannot be derived from the least action principle. Nonetheless, it can be viewed as the high-viscosity limit ($\gamma\dot{W} \gg \ddot{W}$) of the second-order differential equation[2] [66,67]

$$\ddot{W} + \gamma\dot{W} + \frac{\partial C}{\partial W} = 0. \tag{5}$$

This equation resembles the equation of motion for a scalar field $\phi$ in curved spacetime,

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\partial V}{\partial \phi} = 0, \tag{6}$$

when we identify $W$ with $\phi$, $\gamma$ with $3H$, and $C$ with $V$.

Because $\gamma$ is constant in typical training, Eq. (5) suggests a connection between neural network training and dynamics in dS space. Motivated by this analogy, KS proposed that the time evolution of the network output $Z$ obeys

$$\ddot{Z} - \beta\dot{Z} + \Theta\frac{\partial C}{\partial Z} = 0, \tag{7}$$

where $\beta$ is related to $\gamma$ and $\Theta$ denotes the NTK. Because $\Theta$ becomes constant in a certain limit, KS argued that Eq. (7) establishes a duality between the time evolution of $Z$ and cosmological dynamics [5].

To discuss further, KS exploited the EFT approach[3] and conducted a simulation analysis. EFT is indeed a highly promising methodology, but the fundamental degrees of freedom that actually change during training are the neural network parameters, and developing a theory that directly addresses them is worth pursuing. In particular, the analysis by KS rests on the assumption that the empirical NTK [68–70] accurately approximates network dynamics during training. While the NTK captures how parameter variations influence the output, it becomes constant only in restricted cases. Consequently, a more careful examination of how parameters affect the output is required, and constructing a theoretical framework that focuses on the parameters themselves is especially meaningful.

In the following section, we adopt a more fundamental approach, introducing a field-theoretic description of neural network parameters that begins with their microscopic training dynamics. This framework will also reveal a connection to cosmological dynamics, akin to the one explored by KS [5].

*Synaptic field theory.* While the original equation of motion for the synaptic weights (4) is not derived from the least action principle, Eq. (5) can be obtained from the action

$$S = \int dt e^{\gamma t}\left[\frac{1}{2}\dot{W}^2 - C\right], \tag{8}$$

where $\frac{1}{2}\dot{W}^2$ denotes $\sum_{i,j,m}\frac{1}{2}\dot{W}_{ij}^{(m)2}$. The terms inside the brackets resemble a matter Lagrangian, as they follow the typical kinetic minus potential structure. If a suitable continuum limit of Eq. (8) exists, or if the (matter) Lagrangian can be recast as an integral of a Lagrangian density,

$$L[W(t)] = \frac{1}{2}\dot{W}^2 - C = \int d^d x \mathcal{L}[w(t,\mathbf{x})], \tag{9}$$

then the action can be written as

[2]This equation can also be interpreted as describing a neural network with momentum [5,64,65].

[3]Likewise, several studies have sought to develop EFT frameworks for deep learning [5,37,38,50].

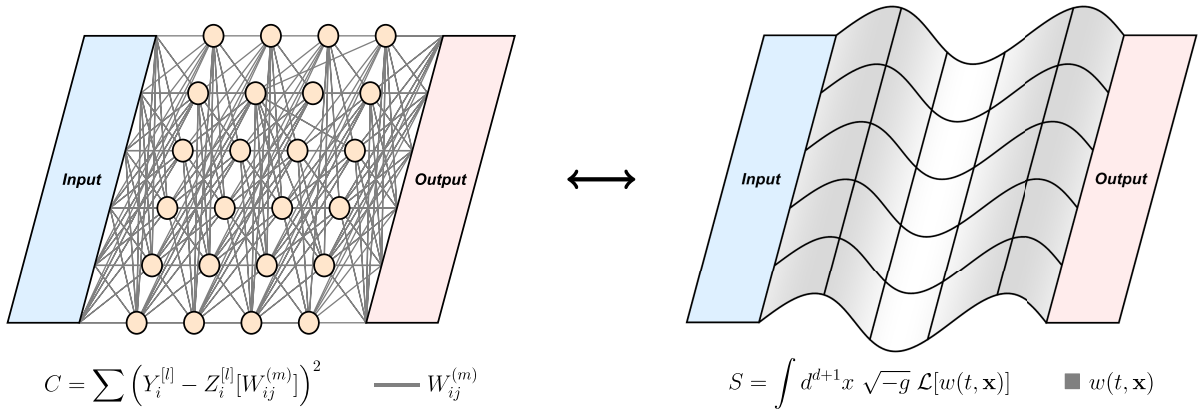



FIG. 1. Visualization of synaptic field theory. We introduce synaptic field theory by interpreting the collective behavior of synaptic weights and biases as a field $w(t,\boldsymbol{x})$ interacting with external sources driven by training datasets. The space of the synaptic field is defined by indices representing neuron positions and their synaptic connections, providing a structured framework to describe neural network dynamics in a field-theoretic manner.

$$S = \int d^{d+1}x e^{\gamma t} \mathcal{L}[w(t,\mathbf{x})]. \tag{10}$$

Identifying $e^{\gamma t} = \sqrt{-g}$ reveals a connection between neural networks and a field theory in curved spacetime. The exponential factor $\sqrt{-g}$ appears in the universe dominated by the cosmological constant. In $(d+1)$-dimensional dS spacetime, the Hubble parameter $H$ becomes a positive constant and the metric tensor takes the form $\sqrt{-g} = e^{dHt}$. Matching exponents identifies constant Hubble parameter $H = \gamma/d$ and indicates the connection to dS space.

Notably, the cost function in Eq. (2) sums over synaptic weight and bias indices. To construct a continuum theory, it is natural to take the continuum limit of these indices, replacing the summation with an integral. We call this theory, defined on the space based on these indices, "synaptic field theory." Figure 1 schematically illustrates this.

If the activation function admits an infinite-series expansion (as the sigmoid does), the cost function can likewise be expressed as an infinite series in the weights

$$C = \sum J^{(m_1)}_{1 i_1 j_1} W^{(m_1)}_{i_1 j_1} + \sum J^{(m_1 m_2)}_{2 i_1 j_1 i_2 j_2} W^{(m_1)}_{i_1 j_1} W^{(m_2)}_{i_2 j_2} + \cdots. \tag{11}$$

The coefficients $J^{(m_1)}_{1 i_1 j_1}$ and $J^{(m_1 m_2)}_{2 i_1 j_1 i_2 j_2}$ depend on the data set, and the expression is defined up to an additive constant. Repeated indices are summed.

Taking the continuum limit—replacing discrete indices by three-dimensional spatial coordinates—yields

$$L \supset \int d^3\mathbf{x} J_1(\mathbf{x}) w(t,\mathbf{x}) + \int d^3\mathbf{x} d^3\mathbf{y} J_2(\mathbf{x},\mathbf{y}) w(t,\mathbf{x}) w(t,\mathbf{y}) + \cdots. \tag{12}$$

Here, $J_1$ and $J_2$ act as external sources determined by the training examples. Table I summarizes how the neural network components map onto elements of the synaptic field theory.

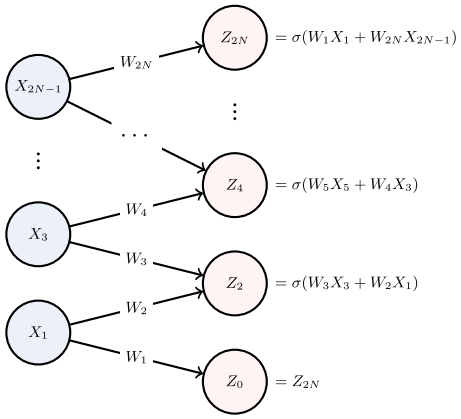
In this approach, it is difficult to endow the synaptic field theory with desirable properties such as locality. The terms in Eq. (12) are inherently nonlocal, and conventional field theory rarely treats genuinely nonlocal theories. Consequently, this nonlocality complicates efforts to formulate a field-theoretic description of neural networks in dS space.

To find the local Lagrangian, it is important to note that the nonlocality is related to the architecture and indexing convention of the parameters. Equation (12) is obtained by naively extending a neural network with a typical indexing scheme. Since these indices are mapped to spatial coordinates in the continuum limit, the architecture and indexing convention of the neural network determine the spatial geometry of the synaptic field theory. It is unclear whether typical architectures and indexing conventions yield geometries that are useful for analysis.

In other words, by designing neural networks with specific architectures and assigning appropriate indices, one may obtain a spatial geometry that is easier to analyze or even admits a local action. In the following discussion,

TABLE I. A dictionary relating neural network components to the synaptic field is presented.

| Neural network | Synaptic field theory |
|---|---|
| Weight $W^{(m)}_{ij}$ | Field $w(t,x)$ |
| Training examples $(X,Y)$ | External sources $J, K, \cdots$ |
| Indices $i$, $j$, and $m$ | Space $x$ |
| Training step $T$ | Time $t$ |
| Cost function $C$ | Lagrangian $L$ |
| Step size $\eta$ | Hubble parameter $H$ |




$$\longleftrightarrow \quad L = \int dx\, \sqrt{-g} \left[ \frac{1}{2}[\partial_t w(t,x)]^2 - \frac{1}{2}K(x)[\partial_x w(t,x)]^2 - \frac{1}{2}J(x)w(t,x)^2 \right]$$

FIG. 2. Neural network architecture (left) and its continuum limit synaptic field theory (right). This illustrates the correspondence between training examples and external sources, as well as the mapping of discrete indices to spatial coordinates in synaptic field theory.

we examine two examples in which local synaptic field theories are derived by adopting an appropriate indexing convention for neural networks with a very simple architecture.

*Examples.* As a first example, consider the perceptron illustrated in Fig. 2, which has an $N$-component input vector and an $N$-component output vector. The input and output layers are connected by weights such that only adjacent components interact, and no bias terms are included. We impose periodic boundary conditions along the width, identifying the $(N+j)$th neuron with the $j$th neuron and the $(2N+k)$th synapse with the $k$th synapse. The activation function is linear, $\sigma(x) = px + q$.

The cost function of this perceptron is

$$C = \sum_{i,l} \left[ Y_{2i}^{[l]} - \sigma\left( W_{2i} X_{2i-1}^{[l]} + W_{2i+1} X_{2i+1}^{[l]} \right) \right]^2. \tag{13}$$

For notational simplicity, odd-numbered indices label inputs, and even-numbered indices label outputs.[4] Because $C$ is a sum of quadratic polynomials in $W_{2i}$ and $W_{2i+1}$, we can expand and shift the weights to recast $C$, up to constant, as

$$C = \sum_i \left[ \frac{1}{2} K_i (W_{i+1} - W_i)^2 + \frac{1}{2} J_i (W_i)^2 \right], \tag{14}$$

where $J_i$ and $K_i$, are coefficients determined by the training data.

This cost function admits the continuum limit under the following heuristic[5] correspondences:

$$\sum_i \to \int dx,$$
$$J_i, K_i, W_i, \Delta W_i \to J(x), K(x), w(t,x), \partial_x w(t,x), \tag{15}$$

where $\Delta W_i$ denotes $W_{i+1} - W_i$ and $w$, $K$, and $J$ are the continuum analogs of $W_i$, $K_i$, and $J_i$, respectively. The continuum limit suggests that the Lagrangian contains the terms

$$\int dx \left[ \frac{1}{2} K(x)[\partial_x w(t,x)]^2 + \frac{1}{2} J(x) w(t,x)^2 \right]. \tag{16}$$

Consequently, the Lagrangian in the continuum limit takes the form of

$$L = \int dx \frac{1}{2} [(\partial_t w)^2 - K(\partial_x w)^2 - J w^2]. \tag{17}$$

As a second example, we examine a more realistic case that differs from the previous one in two respects. First, we insert an additional hidden layer between the input and output layers. Second, we adopt a quadratic activation function, $\sigma(x) = px^2 + qx + r$. As before, we impose periodic boundary conditions and relabel the synapse indices, as illustrated in Fig. 3.

Because the network depth in this example is small, we need not take a continuum limit in the depth direction, unlike in Eq. (12). Instead, we associate each layer with a distinct field: $W_1$ and $W_2$ correspond to $w_1$ and $w_2$, respectively. When the depth becomes large, however, one can apply the continuum limit to the depth index. In that case, the collection $w_1(t,x), w_2(t,x), \ldots$ can be regarded as a single unified field $w(t,\mathbf{x})$, where $\mathbf{x}$ now includes the coordinate associated with depth.

[4]Typically, weights carry two indices, but restricted connections allow a simplified notation.

[5]To clarify what is meant by the continuum limit, a proper length scale—the lattice spacing—must be introduced. In this correspondence, every occurrence of the lattice spacing in the Lagrangian is assumed to be absorbed into the external source.

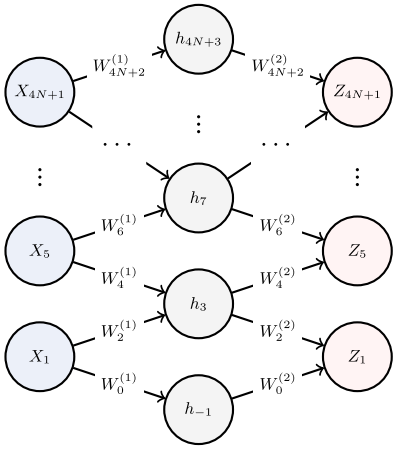



FIG. 3. Neural network architecture for the second example. Periodic boundary conditions are imposed.

The cost function for this neural network is

$$C = \sum_{i,l} \Big[ Y_{4i+1} - W^{(2)}_{4i}\sigma\big(W^{(1)}_{4i}X_{4i+1} + W^{(2)}_{4i-2}X_{4i-3}\big) - W^{(2)}_{4i+2}\sigma\big(W^{(1)}_{4i+2}X_{4i+1} + W^{(1)}_{4i+4}X_{4i+5}\big)\Big]^2, \tag{18}$$

and by following the procedure of Eq. (15), we obtain

$$\begin{aligned} L = \int dx \Big[ & \frac{1}{2}(\partial_t w_1)^2 + \frac{1}{2}(\partial_t w_2)^2 - \frac{1}{2}m^2 w_2^2 \\ & - J_1 - J_2 w_2 - J_3 w_2 w_1 - K_1 w_2 \partial_x w_1 \\ & - K_2 w_2 \partial_x^2 w_2 - K_3 w_2 \partial_x^2 w_1 - K_4 \partial_x w_2 \partial_x w_1 \\ & - K_5 w_1 \partial_x w_2 - K_6 \partial_x^2 w_2 - K_7 w_1 \partial_x^2 w_2 + \cdots \Big]. \end{aligned} \tag{19}$$

Only terms up to quadratic order in $w_1$ and $w_2$, and up to second-order spatial derivatives, are shown. The coupling constants and external sources $m$, $K_i$, and $J_i$ are listed in Table II.

TABLE II. Summary of couplings and external sources in Eq. (19). Here, $a$ and $N_l$ denote the lattice spacing and the size of the training examples, respectively.

| | |
|---|---|
| $m^2 = 8a^{-1}N_l r^2$ | |
| $J_1(x) = a^{-1}\sum_l (Y^{[l]})^2$ | $J_2(x) = 4ra^{-1}\sum_l Y^{[l]}$ |
| $J_3(x) = 8qa^{-1}\sum_l (X^{[l]} + 4a^2\partial_x^2 X^{[l]})Y^{[l]}$ | |
| $K_1(x) = 48aq\sum_l Y^{[l]}\partial_x X^{[l]}$ | $K_2 = 4aN_l r^2$ |
| $K_3(x) = 20qa\sum_l X^{[l]}Y^{[l]}$ | $K_4(x) = 16aq\sum_l X^{[l]}Y^{[l]}$ |
| $K_5(x) = 16aq\sum_l Y^{[l]}\partial_x X^{[l]}$ | $K_6(x) = 2ra\sum_l Y^{[l]}$ |
| $K_7(x) = 4qa\sum_l X^{[l]}Y^{[l]}$ | |

*Discussion and outlooks.* By establishing a concrete bridge between neural networks and field theory, our framework allows neural networks to be analyzed using field-theoretic language. For instance, in Eq. (19), mass terms appear only for the synaptic field of the last layer $w_2$. This is because those parameters are the only ones that can couple to the constant term of the activation function. This feature remains even when additional layers are added in the same way. By analyzing this mass term, the effect of the activation function's value at zero may be investigated. Likewise, further research could explore how specific features of neural networks shape the field-theoretic structure, or vice versa. Several open questions and directions for future work are listed below as examples of such possibilities.

(i) Network structures and properties: As noted above, understanding how the structural features of a neural network are encoded in its synaptic field theory is an intriguing open problem. Elements such as the training data set and the activation function can leave distinctive imprints on the resulting field theory. One could also study training protocols in which the data set or learning rate $\eta$ varies with time; the corresponding theory would then contain time-dependent sources or exhibit expansion histories other than dS—e.g., radiation-dominated evolution. In parallel, developing a synaptic field theory that preserves locality will require a systematic indexing scheme applicable to more general architectures, including networks with complex connectivity, nontrivial activations, and explicit bias terms. For example, if a network's energy or cost function already encodes localitylike structure [71–74], that locality may carry over to the associated field theory.

(ii) Networks with conventional symmetries: Identifying architectures whose continuum limits respect familiar field-theoretic symmetries, such as Poincaré invariance, would be illuminating. Our current constructions generically violate Poincaré invariance because these symmetries are not required in deep learning. Determining the conditions under which they emerge remains a challenging and intriguing problem.

(iii) Statistical-physics perspective: Adding stochastic noise to the gradient-descent dynamics turns the update rule into a Langevin equation, allowing a Gibbs-measure description in which the cost function plays the role of a Hamiltonian. Although noise is absent from our present discussion, incorporating it should naturally link to earlier works [28] and provide a rich statistical field-theoretic framework.

*Acknowledgments.* We thank Krippendorf and Spannowsky for gently pointing out their earlier work, Ref. [5], which first discussed a duality connecting neural network and cosmological dynamics. Regrettably, we were unaware of

this reference when preparing the initial preprint of our manuscript. We thank J. Lee for his helpful comments. This work was supported in part by the National Research Foundation of Korea (Grant No. RS-2024-00352537).

*Data availability.* No data were created or analyzed in this study.

[1] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature (London) **521**, 436 (2015).
[2] F.-L. Fan, J. Xiong, M. Li, and G. Wang, On interpretability of artificial neural networks: A survey, IEEE Trans. Radiat. Plasma Med. Sci. **5**, 741 (2021).
[3] X. Li, H. Xiong, X. Li, X. Wu, X. Zhang, J. Liu, J. Bian, and D. Dou, Interpretable deep learning: Interpretation, interpretability, trustworthiness, and beyond, Knowl. Inf. Syst. **64**, 3197 (2022).
[4] Y. Bahri, J. Kadmon, J. Pennington, S. S. Schoenholz, J. Sohl-Dickstein, and S. Ganguli, Statistical mechanics of deep learning, Annu. Rev. Condens. Matter Phys. **11**, 501 (2020).
[5] S. Krippendorf and M. Spannowsky, A duality connecting neural network and cosmological dynamics, Mach. Learn. Sci. Tech. **3**, 035011 (2022).
[6] K. Hashimoto, S. Sugishita, A. Tanaka, and A. Tomiya, Deep learning and the AdS/CFT correspondence, Phys. Rev. D **98**, 046019 (2018).
[7] K. Hashimoto, S. Sugishita, A. Tanaka, and A. Tomiya, Deep learning and holographic QCD, Phys. Rev. D **98**, 106014 (2018).
[8] K. Hashimoto, AdS/CFT correspondence as a deep Boltzmann machine, Phys. Rev. D **99**, 106017 (2019).
[9] W.-C. Gan and F.-W. Shu, Holography as deep learning, Int. J. Mod. Phys. D **26**, 1743020 (2017).
[10] J. Lee, Y. Bahri, R. Novak, S. S. Schoenholz, J. Pennington, and J. Sohl-Dickstein, Deep neural networks as Gaussian processes, arXiv:1711.00165.
[11] J. M. Antognini, Finite size corrections for neural network Gaussian processes, arXiv:1908.10030.
[12] B. Hanin and M. Nica, Finite depth and width corrections to the neural tangent kernel, arXiv:1909.05989.
[13] J. Huang and H.-T. Yau, Dynamics of deep neural networks and neural tangent hierarchy, arXiv:1909.08156.
[14] E. Dyer and G. Gur-Ari, Asymptotics of wide networks from Feynman diagrams, arXiv:1909.11304.
[15] Y. Sho, Non-Gaussian processes and neural networks at finite widths, arXiv:1910.00019.
[16] G. Naveh, O. Ben David, H. Sompolinsky, and Z. Ringel, Predicting the outputs of finite deep neural networks trained with noisy gradients, Phys. Rev. E **104**, 064301 (2021).
[17] I. Seroussi, G. Naveh, and Z. Ringel, Separation of scales and a thermodynamic description of feature learning in some CNNS, arXiv:2112.15383.
[18] K. Aitken and G. Gur-Ari, On the asymptotics of wide networks with polynomial activations, arXiv:2006.06687.
[19] A. Andreassen and E. Dyer, Asymptotics of wide convolutional neural networks, arXiv:2008.08675.
[20] J. Halverson, A. Maiti, and K. Stoner, Neural networks and quantum field theory, Mach. Learn. Sci. Tech. **2**, 035002 (2021).
[21] J. A. Zavatone-Veth, A. Canatar, B. S. Ruben, and C. Pehlevan, Asymptotics of representation learning in finite Bayesian neural networks, J. Stat. Mech. (2022) 114008.
[22] G. Naveh and Z. Ringel, A self-consistent theory of Gaussian processes captures feature learning effects in finite CNNs, in *Adv. Neural Inf. Process. Syst.* (2021), Vol. 34, p. 21352.
[23] B. Hanin, Random fully connected neural networks as perturbatively solvable hierarchies, arXiv:2204.01058.
[24] S. Yaida, Meta-principled family of hyperparameter scaling strategies, arXiv:2210.04909.
[25] M. Demirtas, J. Halverson, A. Maiti, M. D. Schwartz, and K. Stoner, Neural network field theories: Non-gaussianity, actions, and locality, Mach. Learn. Sci. Technol. **5**, 015002 (2024).
[26] H. Erbin, V. Lahoche, and D. O. Samary, Non-perturbative renormalization for the neural network-QFT correspondence, Mach. Learn. Sci. Tech. **3**, 015027 (2022).
[27] H. Erbin, V. Lahoche, and D. O. Samary, Renormalization in the neural network-quantum field theory correspondence, arXiv:2212.11811.
[28] Z. Ringel, N. Rubin, E. Mor, M. Helias, and I. Seroussi, Applications of statistical field theory in deep learning, arXiv:2502.18553.
[29] M. A. Buice and C. C. Chow, Beyond mean field theory: Statistical field theory for neural networks, J. Stat. Mech. (2013) P03003.
[30] S. Sonoda and N. Murata, Transport analysis of infinitely deep neural network, arXiv:1605.02832.
[31] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks* (Springer Cham, Cham, Switzerland, 2020), Vol. 970.
[32] J. Halverson, Building quantum field theories out of neurons, arXiv:2112.04527.
[33] A. Maiti, K. Stoner, and J. Halverson, Symmetry-via-duality: Invariant neural network densities from parameter-space correlators, arXiv:2106.00694.
[34] J. Halverson, J. Naskar, and J. Tian, Conformal fields from neural networks, arXiv:2409.12222.
[35] K. T. Grosvenor and R. Jefferson, The edge of chaos: Quantum field theory and deep neural networks, SciPost Phys. **12**, 081 (2022).
[36] S. Gukov, J. Halverson, and F. Ruehle, Rigor with machine learning from field theory to the Poincaré conjecture, Nat. Rev. Phys. **6**, 310 (2024).

[37] O. Cohen, O. Malka, and Z. Ringel, Learning curves for overparametrized deep neural networks: A field theory perspective, Phys. Rev. Res. **3**, 023034 (2021).
[38] J. N. Howard, M. S. Klinger, A. Maiti, and A. G. Stapleton, Bayesian RG flow in neural network field theories, SciPost Phys. Core **8**, 027 (2025).
[39] G. Aarts, D. E. Habibi, L. Wang, and K. Zhou, On learning higher-order cumulants in diffusion models, Mach. Learn. Sci. Technol. **6**, 025004 (2025).
[40] M. A. Buice and J. D. Cowan, Field-theoretic approach to fluctuation effects in neural networks, Phys. Rev. E **75**, 051919 (2007).
[41] D. S. Berman and M. S. Klinger, The inverse of exact renormalization group flows as statistical inference, Entropy **26**, 389 (2024).
[42] D. S. Berman, M. S. Klinger, and A. G. Stapleton, Bayesian renormalization, Mach. Learn. Sci. Tech. **4**, 045011 (2023).
[43] J.-W. Lee, Quantum fields as deep learning, J. Korean Phys. Soc. **76**, 684 (2020).
[44] Z.-A. Jia, B. Yi, R. Zhai, Y.-C. Wu, G.-C. Guo, and G.-P. Guo, Quantum neural network states: A brief review of methods and applications, Adv. Quantum Technol. **2**, 1800077 (2019).
[45] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).
[46] D.-L. Deng, X. Li, and S. Das Sarma, Quantum entanglement in neural network states, Phys. Rev. X **7**, 021021 (2017).
[47] D. Bachtis, G. Aarts, and B. Lucini, Quantum field-theoretic machine learning, Phys. Rev. D **103**, 074510 (2021).
[48] W. A. Zúñiga-Galindo, p-adic statistical field theory and deep belief networks, Physica (Amsterdam) **612A**, 128492 (2023).
[49] W. A. Zúñiga-Galindo, C. He, and B. A. Zambrano-Luna, p-Adic statistical field theory and convolutional deep Boltzmann machines, Prog. Theor. Exp. Phys. **2023**, 063A01 (2023).
[50] I. Banta, T. Cai, N. Craig, and Z. Zhang, Structures of neural network effective theories, Phys. Rev. D **109**, 105007 (2024).
[51] V. Vanchurin, Emergent field theories from neural networks, arXiv:2411.08138.
[52] K. Segadlo, B. Epping, A. van Meegen, D. Dahmen, M. Krämer, and M. Helias, Unified field theoretical approach to deep and recurrent neuronal networks, J. Stat. Mech. (2022) 103401.
[53] R. Bondesan and M. Welling, The hintons in your neural network: A quantum field theory view of deep learning, arXiv:2103.04913.
[54] R. M. Neal, Priors for infinite networks, in *Bayesian Learning for Neural Networks* (Springer, New York, 1996), pp. 29–53.
[55] C. K. I. Williams, Computation with infinite neural networks, Neural Comput. **10**, 1203 (1998).
[56] A. G. de G. Matthews, M. Rowland, J. Hron, R. E. Turner, and Z. Ghahramani, Gaussian process behaviour in wide deep neural networks, arXiv:1804.11271.
[57] R. Novak, L. Xiao, J. Lee, Y. Bahri, G. Yang, J. Hron, D. A. Abolafia, J. Pennington, and J. Sohl-Dickstein, Bayesian deep convolutional networks with many channels are Gaussian processes, arXiv:1810.05148.
[58] A. Garriga-Alonso, C. E. Rasmussen, and L. Aitchison, Deep convolutional networks as shallow Gaussian processes, arXiv:1808.05587.
[59] G. Yang, Scaling limits of wide neural networks with weight sharing: Gaussian process behavior, gradient independence, and neural tangent kernel derivation, arXiv:1902.04760.
[60] G. Yang, Tensor Programs I: Wide feedforward or recurrent neural networks of any architecture are Gaussian processes, arXiv:1910.12478.
[61] G. Yang, Tensor Programs II: Neural tangent kernel for any architecture, arXiv:2006.14548.
[62] A. Jacot, F. Gabriel, and C. Hongler, Neural tangent kernel: Convergence and generalization in neural networks, arXiv:1806.07572.
[63] J. Lederer, Activation functions in artificial neural networks: A systematic overview, arXiv:2101.09957.
[64] S. Ruder, An overview of gradient descent optimization algorithms, arXiv:1609.04747.
[65] J. Lee, L. Xiao, S. S. Schoenholz, Y. Bahri, R. Novak, J. Sohl-Dickstein, and J. Pennington, Wide neural networks of any depth evolve as linear models under gradient descent, J. Stat. Mech. (2020) 124002.
[66] G. Parisi, *Statistical Field Theory*, Frontiers in Physics (Addison-Wesley, Redwood City, CA, 1988).
[67] T. L. H. Watkin, A. Rau, and M. Biehl, The statistical mechanics of learning a rule, Rev. Mod. Phys. **65**, 499 (1993).
[68] M. Samarin, V. Roth, and D. Belius, On the empirical neural tangent kernel of standard finite-width convolutional neural network architectures, arXiv:2006.13645.
[69] S. Fort, G. K. Dziugaite, M. Paul, S. Kharaghani, D. M. Roy, and S. Ganguli, Deep learning versus kernel learning: An empirical study of loss landscape geometry and the time evolution of the neural tangent kernel, arXiv:2010.15110.
[70] R. Novak, L. Xiao, J. Hron, J. Lee, A. A. Alemi, J. Sohl-Dickstein, and S. S. Schoenholz, Neural tangents: Fast and easy infinite neural networks in PYTHON, arXiv:1912.02803.
[71] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).
[72] W. Little, The existence of persistent states in the brain, Math. Biosci. **19**, 101 (1974).
[73] P. Smolensky, Information processing in dynamical systems: Foundations of harmony theory, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1—Foundations*, edited by D. E. Rumelhart and J. L. McClelland (MIT Press, Cambridge, MA, 1986), pp. 194–281.
[74] D. Lee, H.-S. Lee, and J. Yi, Dynamic neuron approach to deep neural networks: Decoupling neurons for renormalization group analysis, Phys. Rev. Res. **7**, 023276 (2025).