

Optimal architecture and fundamental bounds in neural network field theory

Zhengkang Zhang *Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA*

(Received 8 July 2026; accepted 10 September 2026; published 30 September 2026)

Neural network field theory (NNFT) represents fields as neural networks and samples field configurations by drawing network parameters from a probability distribution. We identify a previously unexplored architectural freedom in NNFT, parametrized by α , that leaves the infinite-width theory invariant but dramatically affects finite-width errors in the calculation of correlation functions. For a massive scalar field, we show that $\alpha = 0$, corresponding to propagator-weighted neuron momenta and constant neuron amplitudes, is optimal: it minimizes finite-width variance and uniquely removes IR-sensitive corrections in the interacting theory. Even at $\alpha = 0$, relative errors from both bias and variance grow exponentially with distance beyond the correlation length. The bias can be removed by extrapolating to infinite width, which we demonstrate numerically, while the variance imposes a fundamental bound on the achievable signal-to-noise ratio as in lattice field theory. These results chart a path toward developing NNFT into a practical tool for the numerical study of field theories.

DOI: [10.1103/4fzf-tpsv](https://doi.org/10.1103/4fzf-tpsv)

Introduction. Artificial neural networks and Euclidean field theories can both be viewed as collections of random variables. A neural network computes a function of its input, and when the network parameters are drawn from a probability distribution, the output is a random function of the input. A field theory, meanwhile, is a prescription for computing correlation functions of a random field, i.e., a random function of spatial coordinates, $\phi(x)$, where the probability distribution over field configurations is determined by the action $S[\phi]$. Given this structural similarity, a natural question arises: can neural networks with random parameters be used to simulate field theories?

To answer this question, we need to understand how, given a field-theory action, one can construct a neural network architecture and a probability distribution over its parameters such that the resulting random function reproduces the statistics of the field theory, i.e., its correlation functions. This is the core objective of the emerging neural network field-theory (NNFT) program [1–3]. Since its inception, NNFT has been extended in many directions, including conformal field theories [4–6], fermions and supersymmetry [7,8], and string theory [9,10]; see also Refs. [11–18]. In contrast to traditional lattice field theory which discretizes spacetime, NNFT represents the field as a

neural network and samples field configurations by drawing network parameters. This offers potential advantages: the fields live directly in the continuum, and exact Euclidean symmetry can be preserved [19].

The starting point of NNFT is the observation that, by the central limit theorem, a neural network with infinite width (i.e., infinitely many neurons per layer) and independently drawn parameters produces a Gaussian random field, i.e., a free field. This is a field-theoretic restatement of the foundational result in machine learning that infinitely wide neural networks are Gaussian processes [20–25]. To obtain interacting theories, one can break the statistical independence of network parameters while staying in the infinite-width limit [3].

In practice, however, the network width N is finite, which introduces a systematic bias in the form of $1/N$ corrections to the correlation functions. Moreover, the finite ensemble size of Monte Carlo sampling gives rise to statistical variance. A precise characterization of these errors and the limitations they impose is essential for developing NNFT into a practical computational tool.

In this paper, we identify a previously unexplored freedom in NNFT architectures that has dramatic consequences for these errors. We focus on a prototypical massive scalar field theory and parametrize this freedom by α , which corresponds to different ways of distributing the momentum-space integrand between the probability distribution over network parameters and the neuron amplitude. All choices of α yield the same field theory in the $N \rightarrow \infty$ limit, but they give qualitatively different predictions at finite N . Previous work [2,3] adopted

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.

$\alpha = -1$; this particular choice gives rise to (Λ/m) -enhanced bias for correlation functions at distances $r \sim 1/m$ [26], where Λ is the UV cutoff and m is the mass of the scalar field. Here, we systematically extend these results by considering both bias and variance, for all α at all r .

We find that, at $r \gtrsim 1/m$, both bias and variance grow exponentially with mr relative to the signal. This exponential growth represents a qualitatively more severe obstacle than the polynomial enhancement previously identified at $r \sim 1/m$. The bias is a systematic error that can be removed by extrapolating to $N \rightarrow \infty$. The variance, by contrast, represents an irreducible noise floor that persists even as $N \rightarrow \infty$. Nevertheless, we show that choosing $\alpha = 0$ minimizes finite- N contributions above the noise floor. We perform numerical experiments to demonstrate both the removal of bias via extrapolation and α dependence of the variance for the free theory. We also consider the interacting ϕ^4 theory in the perturbative regime and show that $\alpha = 0$ removes IR-sensitive corrections. Together, these results establish $\alpha = 0$ as the optimal architecture for NNFT.

A family of architectures. Consider a one-hidden-layer neural network with cosine activation functions, known as Cos-Net [2,3] (or random Fourier features in machine learning literature [27]):

$$\phi(x; \theta) = \frac{1}{\sqrt{N}} \sum_{i=1}^N w_i^{(2)} \cos(w_i^{(1)} \cdot x + b_i^{(1)}), \quad (1)$$

where $x \in \mathbb{R}^d$ is the input and $\theta = \{w_i^{(1)}, b_i^{(1)}, w_i^{(2)}\}_{i=1}^N$ are the network parameters. When the first-layer weights $w_i^{(1)} \in \mathbb{R}^d$ are drawn from a rotation-invariant distribution and the biases $b_i^{(1)}$ are drawn uniformly from $[-\pi, \pi]$, correlation functions of the network output are Euclidean invariant. We assume the second-layer weights $w_i^{(2)}$ are drawn from a distribution that is independent of $b_i^{(1)}$ but may depend on $w_i^{(1)}$. Defining $a_i \equiv \frac{1}{2} w_i^{(2)} e^{ib_i^{(1)}}$ and $k_i \equiv w_i^{(1)}$, we can rewrite Eq. (1) as a sum of plane waves with random momenta:

$$\phi(x; \theta) = \frac{1}{\sqrt{N}} \sum_{i=1}^N (a_i e^{ik_i \cdot x} + \text{c.c.}). \quad (2)$$

By the central limit theorem, $\phi(x; \theta)$ becomes a Gaussian random field as $N \rightarrow \infty$. The $2n$ -point connected correlators scale as $1/N^{n-1}$, provided the parameters k_i and a_i associated with each neuron are drawn independently from the same distribution. Interactions can be introduced by breaking the independence of network parameters [3], $P(\theta) \propto P_0(\theta) e^{-S_{\text{int}}[\phi(\theta)]}$, or equivalently by reweighting the free-theory ensemble,

$$\langle \phi(x_1) \cdots \phi(x_{2n}) \rangle = \frac{\langle \phi(x_1) \cdots \phi(x_{2n}) e^{-S_{\text{int}}} \rangle_0}{\langle e^{-S_{\text{int}}} \rangle_0}, \quad (3)$$

where $\langle \cdots \rangle_0$ denotes the expectation value with respect to the free-theory distribution $P_0(\theta)$ and we have suppressed the dependence on θ for brevity. In what follows, we first analyze the bias and variance for correlation functions in the free theory and then discuss how these results extend to the interacting theory via Eq. (3).

With independently and identically distributed (i.i.d.) parameters k_i and a_i , the two-point function $G^{(2)}(x_1, x_2) \equiv \langle \phi(x_1) \phi(x_2) \rangle_0$ receives identical contributions from each neuron, yielding

$$G^{(2)}(x_1, x_2) = 2 \int d^d k_i p(k_i) \langle |a_i|^2 \rangle e^{ik_i \cdot (x_1 - x_2)}. \quad (4)$$

Here $p(k_i)$ is the probability distribution over $k_i = w_i^{(1)}$; the average over $b_i^{(1)}$ eliminates the a_i^2 and a_i^{*2} terms, and $\langle |a_i|^2 \rangle$ denotes the average of $|a_i|^2 = (w_i^{(2)})^2/4$ over $w_i^{(2)}$ at fixed $w_i^{(1)}$. No summation over repeated indices is implied throughout this paper. We can choose $p(k_i)$ and $\langle |a_i|^2 \rangle$ to reproduce the propagator in any (reflection-positive) Euclidean-invariant field theory. For a free scalar with mass m , we need

$$p(k_i) \langle |a_i|^2 \rangle = \frac{1}{(2\pi)^d} \frac{f_\Lambda(k_i^2)}{k_i^2 + m^2}, \quad (5)$$

where $f_\Lambda(k_i^2)$ is a UV regulator that satisfies $f_\Lambda(0) = 1$, $f_\Lambda(\infty) = 0$; e.g., one can take $f_\Lambda(k_i^2) = \Theta(\Lambda^2 - k_i^2)$ or $f_\Lambda(k_i^2) = e^{-k_i^2/2\Lambda^2}$.

Equation (5) does not uniquely determine $p(k_i)$ and $\langle |a_i|^2 \rangle$ but only their product [28]. We consider a family of choices parametrized by α :

$$p(k_i) = \frac{1}{(2\pi)^d \Omega_\alpha} \frac{f_\Lambda(k_i^2)}{(k_i^2 + m^2)^{\alpha+1}}, \quad (6)$$

$$\langle |a_i|^2 \rangle = \frac{\Omega_\alpha}{2} (k_i^2 + m^2)^\alpha, \quad (7)$$

where Ω_α is a normalization factor defined as

$$\Omega_\alpha \equiv \int \frac{d^d k}{(2\pi)^d} \frac{f_\Lambda(k^2)}{(k^2 + m^2)^{\alpha+1}}. \quad (8)$$

All choices of α yield the same two-point function, and hence the same $N \rightarrow \infty$ theory, but they give different predictions for higher-point correlation functions at finite N as we will see shortly.

Finite-width bias. Turning to higher-point correlation functions, we define $\phi_i(x) \equiv a_i e^{ik_i \cdot x} + \text{c.c.}$ and denote

the $2n$ -point functions of ϕ and ϕ_i as $G^{(2n)}$ and $G_i^{(2n)}$, respectively. For i.i.d. neurons, we can express $G^{(2n)}$ in terms of $G_i^{(2n')}$ with $n' \leq n$; for example,

$$G^{(4)}(x_1, \dots, x_4) = \left(1 - \frac{1}{N}\right) (G_i^{(2)}(x_1, x_2) G_i^{(2)}(x_3, x_4) + \dots)_{3 \text{ terms}} + \frac{1}{N} G_i^{(4)}(x_1, \dots, x_4), \quad (9)$$

$$G^{(6)}(x_1, \dots, x_6) = \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) (G_i^{(2)}(x_1, x_2) G_i^{(2)}(x_3, x_4) G_i^{(2)}(x_5, x_6) + \dots)_{15 \text{ terms}} + \frac{1}{N} \left(1 - \frac{1}{N}\right) (G_i^{(4)}(x_1, \dots, x_4) G_i^{(2)}(x_5, x_6) + \dots)_{15 \text{ terms}} + \frac{1}{N^2} G_i^{(6)}(x_1, \dots, x_6), \quad (10)$$

where “...” denotes all distinct permutations of the coordinates; see Sec. S1 in Supplemental Material [29] for details. In each equation, the first term is the free-theory prediction rescaled by a factor of $1 + \mathcal{O}(1/N)$, while the remaining terms constitute additional contributions to finite- N bias. The bias from these additional contributions relative to the free-theory prediction is, schematically, controlled by the ratios

$$\frac{1}{N^{n-1}} \frac{G_i^{(2n)}}{[G_i^{(2)}]^n} \equiv \kappa_n, \quad (11)$$

with $n \geq 2$. To determine these ratios, we note that

$$G_i^{(2n)}(x_1, \dots, x_{2n}) = \frac{1}{2} \left(\frac{\Omega_\alpha}{2}\right)^{n-1} \int \frac{d^d k_i}{(2\pi)^d} \frac{\langle |a_i|^{2n} \rangle f_\Lambda(k_i^2)}{\langle |a_i|^2 \rangle^n k_i^2 + m^2} \times (k_i^2 + m^2)^{\alpha(n-1)} \sum_P e^{ik_i \cdot (x_{\text{in}} - x_{\text{out}})}, \quad (12)$$

where the sum runs over all $\binom{2n}{n}$ balanced partitions P of the $2n$ coordinates into “in” and “out” sets, with $x_{\text{in}} \equiv \sum_{j \in \text{in}} x_j$ and $x_{\text{out}} \equiv \sum_{j \in \text{out}} x_j$. In what follows, we take $|a_i| = \sqrt{\Omega_\alpha/2} (k_i^2 + m^2)^{\alpha/2}$ to be deterministic, so that $\langle |a_i|^{2n} \rangle = \langle |a_i|^2 \rangle^n$ [30].

We consider three cases in turn.

- (i) For $\alpha = 0$, Eq. (12) reduces to a sum of propagators, $G_i^{(2n)} \propto \sum_P G^{(2)}(x_{\text{in}} - x_{\text{out}})$, each falling off at large distances as $e^{-m|x_{\text{in}} - x_{\text{out}}|}$. By the triangle inequality, $|x_{\text{in}} - x_{\text{out}}| \leq |x_{j_1} - x_{j_2}| + \dots + |x_{j_{2n-1}} - x_{j_{2n}}|$, where $j_1, \dots, j_{2n}$ is any permutation of $1, \dots, 2n$. Therefore,

$$\kappa_n \sim \frac{e^{cmr}}{N^{n-1}}, \quad (13)$$

where $c > 0$ is an $\mathcal{O}(1)$ constant and r is the characteristic separation between the points—the relative bias grows exponentially with mr at $r \gtrsim 1/m$. Additional enhancements arise from the $(\Omega_\alpha)^{n-1}$ prefactor in Eq. (12), as $\Omega_0 \sim \log(\Lambda/m)$ for $d = 2$ and $\Omega_0 \sim (\Lambda/m)^{d-2}$ for $d \geq 3$.

- (ii) For $\alpha < 0$, the integrand in Eq. (12) is proportional to $1/(k^2 + m^2)^{1+|\alpha|(n-1)}$, which can be written as derivatives of $1/(k^2 + m^2)$ with respect to m^2 . Taking m^2 derivatives does not change the exponential falloff of the integral, so Eq. (13) still holds, with additional enhancements from $\Omega_\alpha \sim (\Lambda/m)^{d-2+2|\alpha|}$.
- (iii) For $\alpha > 0$, it is possible to choose the UV regulator f_Λ such that the integral is exponentially suppressed at $r \gg 1/\Lambda$; e.g., $f_\Lambda(k^2) = e^{-k^2/2\Lambda^2}$ gives $G_i^{(2n)} \sim e^{-\Lambda^2|x_{\text{in}} - x_{\text{out}}|^2/2}$. However, we do not expect the suppression to persist in the interacting theory, because integrating over internal coordinates of interaction vertices will pick up contributions from regions where $|x_{\text{in}} - x_{\text{out}}| \lesssim 1/\Lambda$.

Noise. The finite- N bias discussed above is systematic and can be removed by repeating the calculation at several values of N and extrapolating to $N \rightarrow \infty$. We now turn to the variance component of the error, which unlike the bias cannot be extrapolated away.

Generally, for any Monte Carlo approach that estimates correlation functions by averaging over M samples of the random function ϕ ,

$$\hat{G}^{(2n)}(x_1, \dots, x_{2n}) \equiv \frac{1}{M} \sum_{m=1}^M \phi^{(m)}(x_1) \dots \phi^{(m)}(x_{2n}), \quad (14)$$

the variance of the estimator $\hat{G}^{(2n)}$ is given by

$$\begin{aligned} \text{Var}[\hat{G}^{(2n)}(x_1, \dots, x_{2n})] \\ = \frac{1}{M} [\langle \phi(x_1)^2 \dots \phi(x_{2n})^2 \rangle - \langle \phi(x_1) \dots \phi(x_{2n}) \rangle^2]. \end{aligned} \quad (15)$$

The first term is a $4n$ -point function with coincident points, which does not fall off exponentially and dominates the variance at large distances. In NNFT with i.i.d. neurons, we can express this $4n$ -point function in terms of single-neuron correlation functions similar to Eqs. (9) and (10); see Sec. S1 in Supplemental Material [29]. In the $N \rightarrow \infty$ limit,

$$\text{Var}[\hat{G}^{(2n)}(x_1, \dots, x_{2n})] \simeq \frac{1}{M} \prod_{j=1}^{2n} G_i^{(2)}(x_j, x_j) = \frac{(\Omega_0)^{2n}}{M}. \quad (16)$$

Equation (16) defines an α -independent noise floor and implies a signal-to-noise ratio (SNR) that falls off exponentially with mr at large distances:

$$\text{SNR} \sim \sqrt{M} \frac{G^{(2n)}}{(\Omega_0)^n} \sim \sqrt{M} e^{-nmr}. \quad (17)$$

To achieve $\mathcal{O}(1)$ SNR, we therefore need $M \gtrsim e^{2nmr}$. This is the same signal-to-noise problem that was first analyzed by Parisi and Lepage in the context of lattice field theory [31,32]. The architecture-independent noise floor represents a fundamental cost of the NNFT approach.

On the other hand, at finite N , there are additional contributions to the variance, making Eq. (17) an upper bound on the achievable SNR. We find that, similar to the bias, the size of finite- N contributions relative to the noise floor is controlled by the ratios κ_n defined in Eq. (11), now evaluated at an even number of points that are pairwise coincident to yield

$$\kappa_n^{\text{noise}} = \left(\frac{1}{2N} \right)^{n-1} \frac{(\Omega_\alpha)^{n-1} \Omega_{-\alpha(n-1)}}{(\Omega_0)^n}. \quad (18)$$

From the definition of Ω_α in Eq. (8), it is straightforward to show that Eq. (18) is minimized at $\alpha = 0$ for all $n \geq 2$; see Sec. S2 in Supplemental Material [29] for details. For $\alpha = 0$, $\kappa_n^{\text{noise}} = 1/(2N)^{n-1} \ll 1$ and we nearly saturate the bound in Eq. (17); for either $\alpha < 0$ or $\alpha > 0$, it is possible for the finite- N contributions to dominate over the noise floor due to enhancements by powers of Λ/m .

Numerical results. To test the analytic predictions above, we evaluate $G^{(4)}$ for 50 randomly generated four-point configurations in $d = 2$ using an ensemble of $M = 10^8$ Cos-Nets at each width $N = 30, 50, 100$. We choose a Gaussian regulator, $f_\Lambda(k^2) = e^{-k^2/2\Lambda^2}$, with $\Lambda = 100m$. For each configuration, we rescale the coordinates of the four points by a common factor to achieve three characteristic separations $r = 0.3 \text{ m}^{-1}, m^{-1}$, and 3 m^{-1} , where r is defined as the average of the six pairwise distances between the four points. Figure 1 shows the ratio of the NNFT result to the free-theory prediction as a function of $1/N$ for different values of α and r for the first five configurations; the same qualitative behavior is observed for the other 45 configurations. A weighted linear fit in $1/N$ is extrapolated to $1/N = 0$. The simulation and plotting code was written by Anthropic's Claude Opus 4.5, 4.6 and 4.7, accessed through the Claude Code interface and directed in natural language by the author; its correctness was verified by direct inspection.

The contrast between different α is striking. For $\alpha = 0$ (center column), the free-theory prediction is reproduced within much smaller uncertainties than the other two cases upon extrapolation to $1/N = 0$, confirming the optimality of this choice. For $\alpha = -1$ (left column), the finite- N ratios deviate from unity by orders of magnitude, reflecting (Λ/m) -enhanced bias relative to the $\alpha = 0$ case. For $\alpha = 1$ (right column), the ratios are consistent with unity at all N

within large error bars, confirming the suppressed bias but enhanced variance discussed above. In all cases, we see a substantial increase in both finite- N bias and variance with r , consistent with the analytically predicted exponential growth. Across all 50 configurations and all (α, r) combinations, the extrapolated intercepts are consistent with unity within uncertainties, validating the N -extrapolation approach for removing bias. The irreducible noise floor, on the other hand, presents an obstacle to achieving high precision at $r \gtrsim m^{-1}$. We have also performed numerical experiments for $d = 1, 3, 4$ and observed the same qualitative behavior as in $d = 2$.

Interacting theory. We can extend the bias and variance analysis to the interacting theory via Eq. (3). Consider ϕ^4 theory as an example, with $S_{\text{int}} = \frac{\lambda}{4!} \int d^d x \phi(x)^4$. To analytically probe the α dependence, we focus on the perturbative regime and expand the right-hand side of Eq. (3) in λ . The leading correction to the two-point function is given by

$$\langle \phi(x_1) \phi(x_2) \rangle|_{\mathcal{O}(\lambda)} = -\frac{\lambda}{4!} \int d^d x [G^{(6)}(x_1, x_2, x, x, x, x) - G^{(2)}(x_1, x_2) G^{(4)}(x, x, x, x)], \quad (19)$$

where $G^{(2)}$, $G^{(4)}$ and $G^{(6)}$ are the free-theory correlation functions. Substituting in Eqs. (9) and (10) and integrating over the vertex position x , we find that a subset of the finite- N corrections have integrands independent of x , yielding contributions proportional to the spatial volume [26]. At $\mathcal{O}(1/N)$, these are proportional to

$$\Omega_\alpha \Omega_{-\alpha}(x_1 - x_2) - \Omega_0 \Omega_0(x_1 - x_2), \quad (20)$$

where $\Omega_\alpha(r) \equiv \int \frac{d^d k}{(2\pi)^d} \frac{f_\Lambda(k^2)}{(k^2 + m^2)^{\alpha+1}} e^{ik \cdot r}$ generalizes Ω_α in Eq. (8) to nonzero r . These and analogous contributions at $\mathcal{O}(1/N^2)$ vanish if and only if $\alpha = 0$; see Sec. S3 in Supplemental Material [29] for details. This provides an independent argument for the optimality of $\alpha = 0$, complementing the free-theory analysis. Even for $\alpha = 0$, the remaining finite- N corrections include contributions that fall off more slowly than the field-theory prediction, and the relative bias grows exponentially with mr as in the free theory. We have verified that both the cancellation of IR divergences at $\alpha = 0$ and the exponential growth of relative bias persist for the four-point function up to $\mathcal{O}(\lambda^2)$; a detailed analysis will be presented in a forthcoming paper.

Summary and outlook. In this paper, we have identified a previously unexplored architectural freedom—parametrized by α —in NNFT, a recently proposed approach for simulating field theories using neural

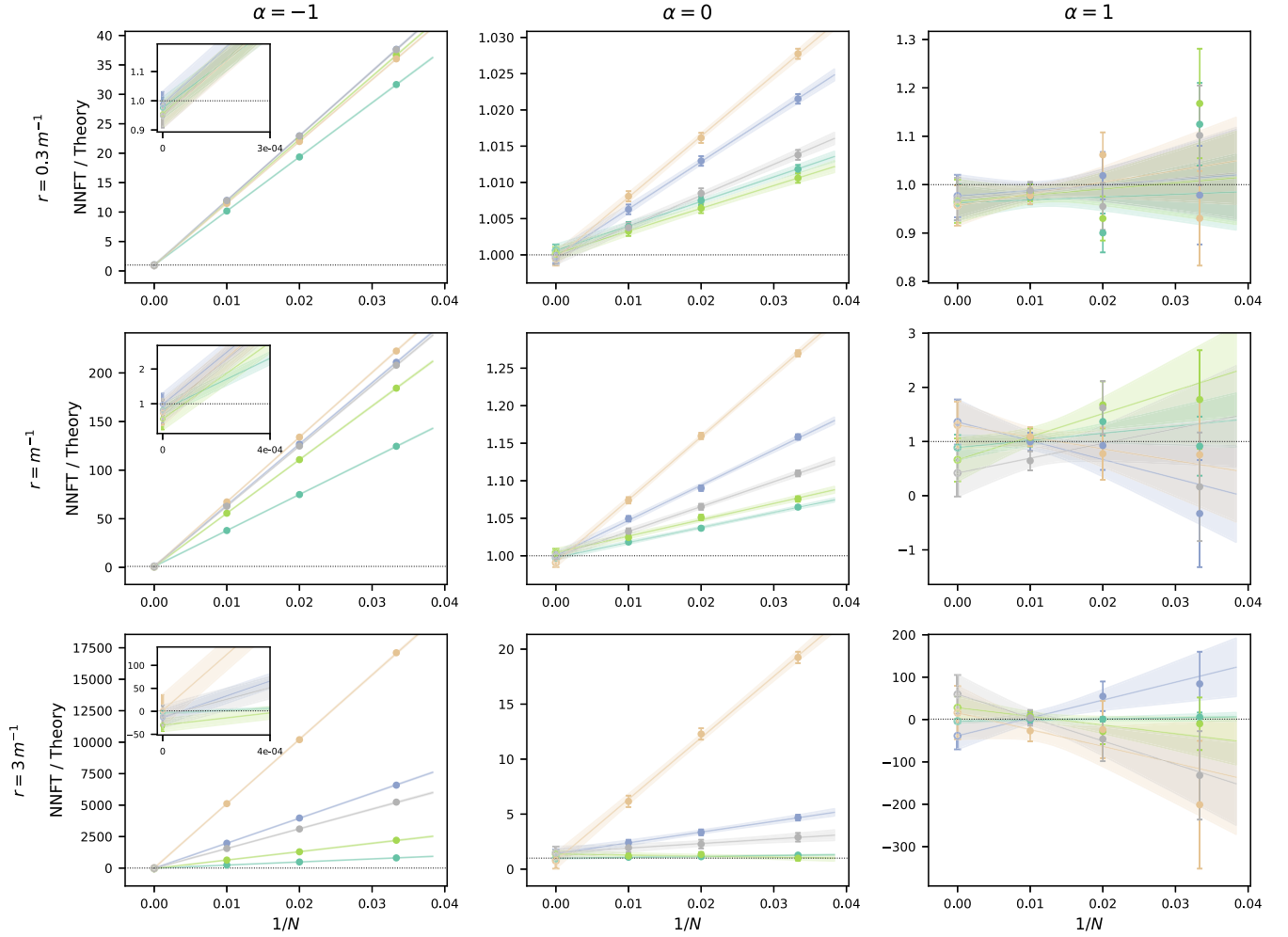


FIG. 1. Ratio of the NNFT four-point function to the free-theory prediction as a function of $1/N$, for five random configurations. Columns correspond to $\alpha = -1, 0, 1$; rows to characteristic separations $r = 0.3m^{-1}, m^{-1}, 3m^{-1}$. Shaded bands indicate weighted linear fits and open circles mark the extrapolated intercepts at $1/N = 0$. Dotted line marks $\text{NNFT}/\text{theory} = 1$. While previous work adopted $\alpha = -1$, we find that $\alpha = 0$ is optimal in minimizing finite- N errors. The simulations and this figure were produced with code written by Claude Opus 4.5, 4.6 and 4.7 under the author's direction, as described in the text.

network ensembles. While all choices of α yield the same field theory in the limit of infinite network width $N \rightarrow \infty$, they give rise to drastically different finite- N behaviors. We have shown that $\alpha = 0$, corresponding to sampling neuron momenta in direct proportion to the propagator while keeping the neuron amplitudes constant, is optimal for minimizing finite- N errors in the evaluation of correlation functions.

Even for the optimal $\alpha = 0$ architecture, relative errors from both bias and variance grow exponentially with distance r at $r \gtrsim 1/m$ —a qualitatively more severe limitation than the polynomial Λ/m enhancement previously identified at $r \sim 1/m$. The bias can be removed by extrapolating to $N \rightarrow \infty$, as we have demonstrated numerically. The variance, on the other hand, sets an upper bound

on the achievable SNR and imposes an exponential cost for measuring correlation functions at large distances, just as in lattice field theory.

Together, these results establish a practical strategy for turning NNFT from a proof-of-principle construction into a computational tool: adopt the optimal $\alpha = 0$ architecture, extrapolate to $N \rightarrow \infty$ to remove bias, and maximize the ensemble size to suppress the noise floor. That the signal-to-noise problem is shared with lattice field theory suggests that techniques developed in that context, such as variance reduction methods, may be applicable to NNFT as well. The combination of these techniques with the continuum formulation intrinsic to NNFT may ultimately open a new frontier in nonperturbative field-theory computations.

Acknowledgments. Z. Z. is supported by the U.S. National Science Foundation under Grant No. PHY-2412880. We thank Mehmet Demirtas and Ben Michel for facilitating initial access to Claude Max. In addition to writing the code, Claude Opus 4.6 and 4.7 provided useful feedback on the manuscript; the analytic results, the design of the

numerical experiments and the interpretation of their outcome are Z. Z.'s own, and Z. Z. takes full responsibility for the content and accuracy of this paper.

Data availability. The code and data supporting the findings of this article are publicly available [33].

-
- [1] J. Halverson, A. Maiti, and K. Stoner, Neural networks and quantum field theory, *Mach. Learn. Sci. Tech.* **2**, 035002 (2021).
 - [2] J. Halverson, Building quantum field theories out of neurons, [arXiv:2112.04527](#).
 - [3] M. Demirtas, J. Halverson, A. Maiti, M. D. Schwartz, and K. Stoner, Neural network field theories: Non-Gaussianity actions, and locality, *Mach. Learn. Sci. Tech.* **5**, 015002 (2024).
 - [4] J. Halverson, J. Naskar, and J. Tian, Conformal fields from neural networks, *J. High Energy Phys.* **10** (2025) 039.
 - [5] P. Capuzzo, B. Robinson, and B. Suzzoni, Conformal defects in neural network field theories, *J. High Energy Phys.* **05** (2026) 124.
 - [6] B. Robinson, Virasoro symmetry in neural network field theories, [arXiv:2512.24420](#).
 - [7] G. Huang and K. Zhou, The neural networks with tensor weights and emergent fermionic Wick rules in the large-width limit, *Phys. Lett. B* **873**, 140146 (2026).
 - [8] S. Frank, J. Halverson, A. Maiti, and F. Ruehle, Fermions and supersymmetry in neural network field theories, *Mach. Learn. Sci. Tech.* **7**, 045070 (2026).
 - [9] S. Frank and J. Halverson, String theory from infinite width neural networks, [arXiv:2601.06249](#).
 - [10] D. S. Ageev and Y. A. Ageeva, Excited string states and D-branes from infinite width neural networks, [arXiv:2602.10214](#).
 - [11] H. Erbin, V. Lahoche, and D. O. Samary, Non-perturbative renormalization for the neural network-QFT correspondence, *Mach. Learn. Sci. Tech.* **3**, 015027 (2022).
 - [12] H. Erbin, V. Lahoche, and D. O. Samary, Renormalization in the neural network-quantum field theory correspondence, [arXiv:2212.11811](#).
 - [13] 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).
 - [14] J. Halverson, TASI lectures on physics for machine learning, [arXiv:2408.00082](#).
 - [15] C. Ferko and J. Halverson, Quantum mechanics and neural networks, *Mach. Learn. Sci. Tech.* **7**, 015002 (2026).
 - [16] C. Ferko, J. Halverson, and A. Mutchler, Universality of neural network field theory, [arXiv:2601.14453](#).
 - [17] D. S. Ageev and Y. A. Ageeva, Neural network quantum field theory from transformer architectures, [arXiv:2602.10209](#).
 - [18] C. Ferko, J. Halverson, V. Jejjala, and B. Robinson, Topological effects in neural network field theory, [arXiv:2604.02313](#).
 - [19] A. Maiti, K. Stoner, and J. Halverson, Symmetry-via-duality: Invariant neural network densities from parameter-space correlators, [arXiv:2106.00694](#).
 - [20] R. M. Neal, *Bayesian Learning for Neural Networks*, Lecture Notes in Statistics, Vol. 118 (Springer, New York, 1996).
 - [21] C. K. I. Williams, Computing with infinite networks, in *Advances in Neural Information Processing Systems*, Vol. 9, edited by M. Mozer, M. Jordan, and T. Petsche (MIT Press, Cambridge, MA, 1996).
 - [22] J. Lee, Y. Bahri, R. Novak, S. S. Schoenholz, J. Pennington, and J. Sohl-Dickstein, Deep neural networks as gaussian processes, [arXiv:1711.00165](#).
 - [23] A. G. d. G. Matthews, M. Rowland, J. Hron, R. E. Turner, and Z. Ghahramani, Gaussian process behaviour in wide deep neural networks, [arXiv:1804.11271](#).
 - [24] G. Yang, Tensor programs I: Wide feedforward or recurrent neural networks of any architecture are gaussian processes, [arXiv:1910.12478](#).
 - [25] B. Hanin, Random neural networks in the infinite width limit as gaussian processes, [arXiv:2107.01562](#).
 - [26] S. Sen and V. Vaidya, Viability of perturbative expansion for quantum field theories on neurons, *Mach. Learn. Sci. Tech.* **7**, 035044 (2026).
 - [27] A. Rahimi and B. Recht, Random features for large-scale kernel machines, in *Advances in Neural Information Processing Systems*, Vol. 20, edited by J. Platt, D. Koller, Y. Singer, and S. Roweis (Curran Associates, Inc., Red Hook, NY, 2007), pp. 1177–1184.
 - [28] This freedom was noted in passing in Ref. [9] in the context of the 2D free boson.
 - [29] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/4fzf-tpsv> for the derivation of the correlation functions for i.i.d. neurons, the minimization of κ_n^{noise} , and the full $\mathcal{O}(\lambda)$ result for the two-point function in ϕ^4 theory.
 - [30] Sampling $|a_i|$ from a Gaussian as in Refs. [2,3] instead gives $\langle |a_i|^{2n} \rangle = (2n-1)!! \langle |a_i|^2 \rangle^n$, which amplifies the bias and variance by numerical prefactors but does not change the parametric dependence on α .
 - [31] G. Parisi, The strategy for computing the hadronic mass spectrum, *Phys. Rep.* **103**, 203 (1984).
 - [32] G. P. Lepage, The analysis of algorithms for lattice field theory, in *Theoretical Advanced Study Institute in Elementary Particle Physics* (World Scientific, Singapore, 1989).
 - [33] Z. Zhang, nnft-alpha: Simulation code and data for Optimal Architecture and Fundamental Bounds in Neural Network Field Theory, <https://github.com/zzkevin2019/nnft-alpha> (2026).