Universality, Topology, and Symmetry in Neural Network Field Theory

Christian Ferko

Northeastern University and IAIFI, the
Institute for Artificial Intelligence and Fundamental Interactions

The 11th LITP Spring Symposium: Theoretical Physics and AI
Leinweber Institute for Theoretical Physics – University of Michigan
May 18, 2026

Recap: neural network field theory.

Thanks to  for the great talk and for nicely teeing up this one!

To review, the idea of NN-FT is to replace the familiar path integral with an ordinary integral over parameters or "latent variables":

$$\begin{aligned}\langle \phi(x_1) \ldots \phi(x_n) \rangle &= \int \mathcal{D}\phi \, e^{-S[\phi]} \phi(x_1) \ldots \phi(x_n) \\ &\longrightarrow \int d\theta \, p(\theta) \phi_\theta(x_1) \ldots \phi_\theta(x_n) \, .\end{aligned}$$

The defining data of a NN-FT is therefore the **architecture** $\phi_\theta(x)$ as well as the **parameter density** $p(\theta)$. This allows us to compute correlators.

This talk is about how **general** and **useful** this framework is. Which features of conventional QFT can be described by this formalism?

Roadmap.

Goal: explore aspects of conventional quantum field theory, such as interactions, topology, and symmetry, that can be captured in NN-FT.

I will give a rapid-fire sample of a few recent results:

☑ Part 1: Introduction and motivation.

☐ Part 2: Universality and Liouville from NN-FT.

☐ Part 3: Topological effects in NN-FT.

☐ Part 4: Symmetries and anomalies in NN-FT.

☐ Part 5: Summary and future directions.

Part 2: Universality and Liouville from NN-FT.

Existence and universality.

The first natural question is under what conditions a given QFT could even be represented by a NN. Fortunately, there are broad existence results.

1. In one Euclidean dimension, under mild assumptions, *any* quantum mechanical theory admits a NN representation [Ferko, Halverson '25].
2. In higher dimensions, any probability measure over tempered distributions $\mathcal{S}'(\mathbb{R}^d)$ admits a neural-network description with countably many parameters [Ferko, Halverson, Mutchler '26].

The proof of the latter is somewhat technical, but boils down to an application of the Borel isomorphism theorem in measure theory.

Commuting diagram for NN-QFT.

Imagine taking a continuum limit of a kinetic term with spacing a

$$d^d x \, \partial^\mu \phi \partial_\mu \phi \longrightarrow a^d \left(\frac{\phi(x+an) - \phi(x)}{a} \right)^2 ,$$

where n is a lattice vector. This term remains finite as $a \to 0$ if $\phi(x+an) - \phi(x) \sim a^{1-d/2}$. In higher d, objects with this scaling are subspaces of $\mathcal{S}'(\mathbb{R}^d)$. The universality statement is that *any* measure over such objects can be represented by a NN with countably many parameters.

$$\begin{tikzcd}
(\Omega, \mathcal{F}) \arrow[r, "\phi"] \arrow[d, "\Theta"'] & (\mathcal{S}'(\mathbb{R}^d), \mathcal{B}(\mathcal{S}'(\mathbb{R}^d))) \\
(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R})^{\mathbb{N}}) \arrow[ur, dashed, "\phi_\theta"'] &
\end{tikzcd}$$

Liouville.

The universality result is non-constructive, so one might ask whether we can actually realize NN-FTs in practice. Consider Liouville theory,

$$S = \int_{\Sigma} d^2x \sqrt{g} \left(\frac{1}{4\pi} \partial_\mu \phi \partial^\mu \phi + \frac{Q}{4\pi} R \phi + \mu e^{2b\phi} \right) ,$$

where μ, b are parameters and $Q = b + \frac{1}{b}$.

The key primary operators in Liouville theory are vertex operators

$$V_\alpha =: e^{2\alpha\phi(x)} : , \qquad \Delta_\alpha = \alpha(Q - \alpha) .$$

The three-point function of such operators is given by a closed-form expression called the DOZZ formula. Can we match DOZZ with NN-FT?

Start Gaussian, then deform.

Following the strategy of [Demirtas, Halverson, Maiti, Schwartz, Stoner '23], we

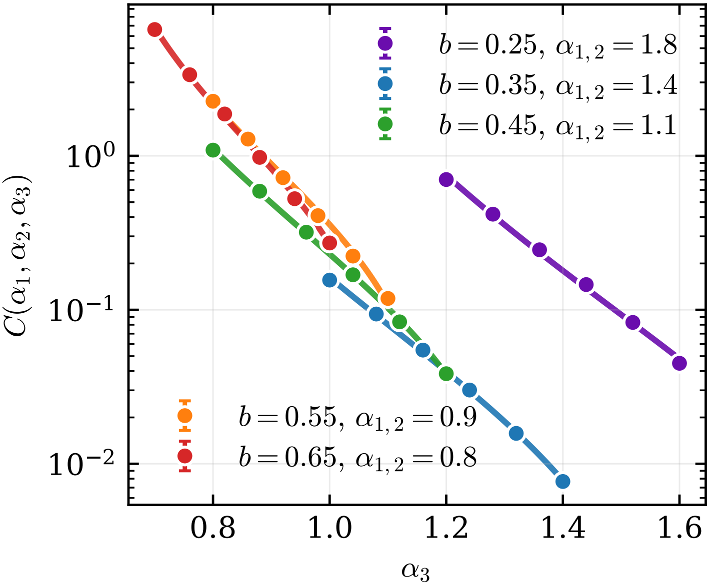
1. First split the field $\phi = c + X$ where c is the zero mode.
2. Begin with a *free field* architecture for X,

   $$X = \sum_{\ell=1}^{L} \sum_{m=-\ell}^{\ell} a_{\ell,m} Y_{\ell,m} \,, \quad a_{\ell,m} \sim \mathcal{N}\left(0, \sigma^2 = \frac{4\pi}{\ell(\ell+1)}\right) .$$

   This is a single-layer neural network with independent Gaussian parameters, and "neurons" which are real spherical harmonics.

   For finite L, this is a function. As $L \to \infty$, it does not converge to a function but a genuine Schwartz distribution. One must therefore be careful about mollifying (smearing) and normal-ordering.
3. *Deform* the parameter density to incorporate interactions. This can be done by breaking independence: we choose $P(c \mid a_{\ell,m})$, so the zero mode distribution depends on the draws of the $a_{\ell,m}$.

Numerical match to DOZZ formula.



Part 3: Topological effects in NN-FT.

Compact targets require discrete latent variables.

For non-compact bosons, a neural network can cleanly represent local fluctuations [Halverson '21]. But compact theories contain additional data: **winding sectors, vortices, momentum sectors, defects, ...**

These features cannot be easily encoded in a single neural network architecture. A NN is a single-valued function; fields with, e.g., winding configurations are in a sense multi-valued.

The natural extension is therefore a mixed continuous/discrete ensemble,

$$\langle \mathcal{O} \rangle = \sum_{Q} \int d\theta \, P(\theta, Q) \, \mathcal{O}[\phi_{\theta,Q}] \, ,$$

where θ are continuous parameters and Q are discrete topological labels.

This mirrors ordinary field theory. For instance, local fluctuations in $\mathcal{L}_{\text{YM}} = -\frac{1}{4} \operatorname{tr}(F_{\mu\nu} F^{\mu\nu})$ would not have told you about the existence of other instanton sectors; one must explicitly sum over such sectors.

Case study: Berezinskii-Kosterlitz-Thouless transition.

The BKT transition is a phase transition in the $2d$ XY model, whose continuum version flows in the IR to the theory of a $2d$ scalar.

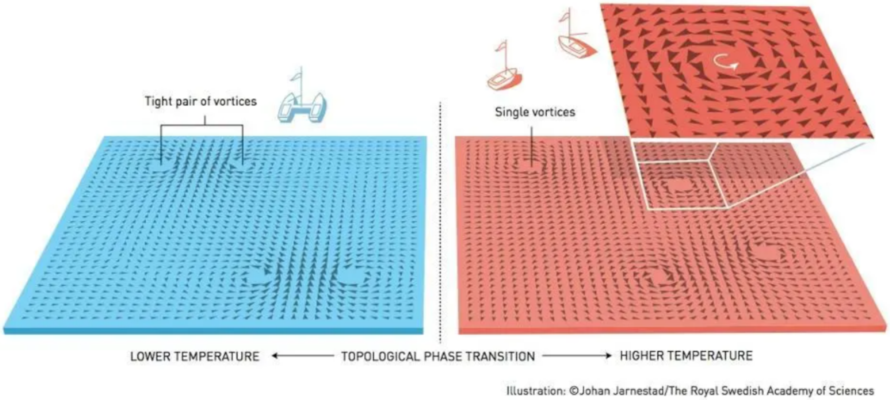


Illustration: ©Johan Jarnestad/The Royal Swedish Academy of Sciences

Low-T is controlled by a "spin-wave" action $S_{\text{sw}} = \frac{K_R}{2} \int d^2x \, (\nabla\theta)^2$, $\theta(x) \sim \theta(x) + 2\pi$; K_R is a "dimensionless stiffness".

Our BKT architecture.

For the compact boson / XY model, the NN-FT field is decomposed as

$$\theta(x) = b\,\theta_{\mathrm{sw}}(x) + \theta_v(x)\,.$$

The two sectors are sampled independently:

1. **Spin waves:** a random Fourier feature network

$$\theta_{\mathrm{sw}}(x) = \frac{A}{\sqrt{N}} \sum_{j=1}^{N} \cos(k_j \cdot x + \gamma_j)$$

   with $k_j = \frac{2\pi}{L} n_j$ and $p(n) \sim 1/|n|^2$ producing the $2d$ logarithmic correlator. Here γ_j is a random phase drawn uniformly on $[0, 2\pi]$.

2. **Vortices:** an explicit singular field

$$\theta_v(x) = \sum_a m_a \arg(x - x_a)\,, \qquad m_a = \pm 1\,,$$

   sampled from a neutral Coulomb gas, where arg is the polar angle.

The low-temperature physics is continuous, while the topological defects are encoded by discrete latent variables.

What the mixed ensemble reproduces in BKT.

The temperature is related to a parameter b by $T = 2\pi\rho_{s,0}b^2$, where $\rho_{s,0}$ is a "bare stiffness", and the phase transition occurs at $b_c = \frac{1}{2}$.

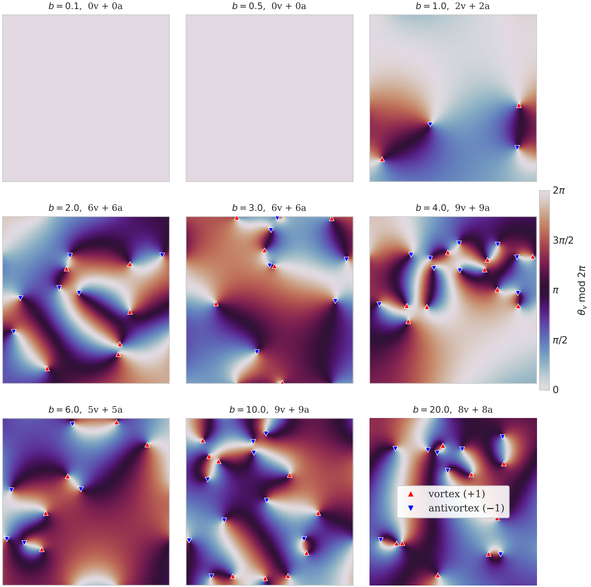
The key behavior that our architecture reproduces is a change in the correlator from pure power-law behavior to an exponentially-suppressed power law above the critical point:

$$G_2(r) \sim \begin{cases} r^{-b^2} \,, & b < b_c \,, \\ r^{-b^2} e^{-r/\xi} \,, & b > b_c \,. \end{cases}$$

Our numerics also reproduce other expected behavior, such as a rapid rise in vortex density above the BKT phrase transition.

Without the discrete vortex sector, the theory never disorders.

Visualization of pure vortex sector.



Part 4: Symmetries and anomalies in NN-FT.

A general framework for symmetry and breaking.

Some symmetries can be easily incorporated into NN-FT; e.g. the uniform phases in the preceding spin wave architecture give translation invariance.

To be more systematic, consider a parameter-space vector field ξ^a and

$$\delta_\xi \phi^\alpha(x) = \sum_a \xi^a \partial_a \phi^\alpha(x) \, .$$

One can derive a parameter-space Schwinger-Dyson equation

$$\sum_{k=1}^{n} \left\langle \phi^{\alpha_1}(x_1) \cdots \delta_\xi \phi^{\alpha_k}(x_k) \cdots \phi^{\alpha_n}(x_n) \right\rangle = - \left\langle B(\theta) \, \phi^{\alpha_1}(x_1) \cdots \phi^{\alpha_n}(x_n) \right\rangle$$

where $B(\theta)$ is called the *breaking function*,

$$B(\theta) := \sum_a \xi^a \, s_a + \sum_a \partial_a \xi^a \, .$$

and $s_a = \partial_a \log \left(p(\theta) \right)$ is called the *score*.

B diagnoses anomalies.

The Schwinger-Dyson equation holds for any ξ_a. In the special case where the variation is a symmetry, one obtains Ward identities.

By definition, a symmetry leaves the partition function $Z[J]$ invariant. The breaking function should therefore also play a role in $\delta Z[J]$, and indeed

$$\delta Z[J] = -\left\langle B(\theta) e^{\int d^d x \sum_\alpha J_\alpha \phi^\alpha_\theta} \right\rangle Z[0] \, .$$

Thus $B(\theta) = 0$ implies that the theory is invariant under an absorbed transformation generated by ξ_a. If a putative symmetry can be recast via such a ξ_a, then the breaking function therefore tests whether it is a genuine symmetry of the theory, or if it is "anomalous".[*]

---

[*] This is not saying that the symmetry leaves the action invariant but transforms the measure; there is no clear separation between classical and quantum in NN-FT.

Example: critical dimension $D = 26$.

As an application of this formalism, we considered the Weyl anomaly of bosonic string theory on a spherical worldsheet.

1. The matter sector X^μ uses a spherical harmonic architecture, similar to Liouville theory before introducing interactions.
2. We engineer the *bc* ghost system using related "spin-weighted spherical harmonics" and Grassmann parameters.
3. Under $g \to e^{2\epsilon} g$, the spherical harmonics and sphere Laplacian change in a way that can be absorbed into a parameter flow ξ_a.
4. The change in the unnormalized log partition function is $\delta_\epsilon \log Z[g] = -\epsilon B$, and we compute B using the preceding formula.

We find that $B = -\frac{D-26}{3}$ vanishes if and only if $D = 26$, as expected.

Part 5: Summary and future directions.

Summary.

We have given an overview of a research direction where a QFT is *defined* as a neural network with an architecture and parameter density.

1. *Any* quantum field theory described by a probability measure over appropriate "field configurations" (i.e. distributions) admits a NN description with countably many parameters.
2. As an example, the Liouville CFT can be engineered as a NN-FT. Our NN-Liouville implementation reproduces DOZZ to $\sim 1\%$.
3. Topological effects, such as winding and vortices, can be naturally incorporated into the NN-FT paradigm through the addition of discrete parameters that label topological sectors.
4. Symmetries and anomalies can be described in NN-FT using a "breaking function" framework. Demanding the vanishing of the breaking function for the Weyl anomaly reproduces the critical dimension $D = 26$ of the bosonic string.

Further research.

There are several potential directions for extending our analysis.

1. Use *training* to construct NN-FT models that reproduce desired features (correlators, energies, etc.) of a “target” QFT.
2. Understand when $d \geq 2$ NN-FT constructions satisfy consistency conditions such as the Osterwalder–Schrader axioms.
3. Construct other interacting $d \geq 2$ field theories using neural networks, e.g. ϕ^4 in $d = 2, 3$ or Yang-Mills in $d = 4$.

Progress on these directions could improve our understanding of functional integration using neural network tools, and perhaps help us develop a clearer picture of what a quantum field theory really is.

Thank you for your attention!