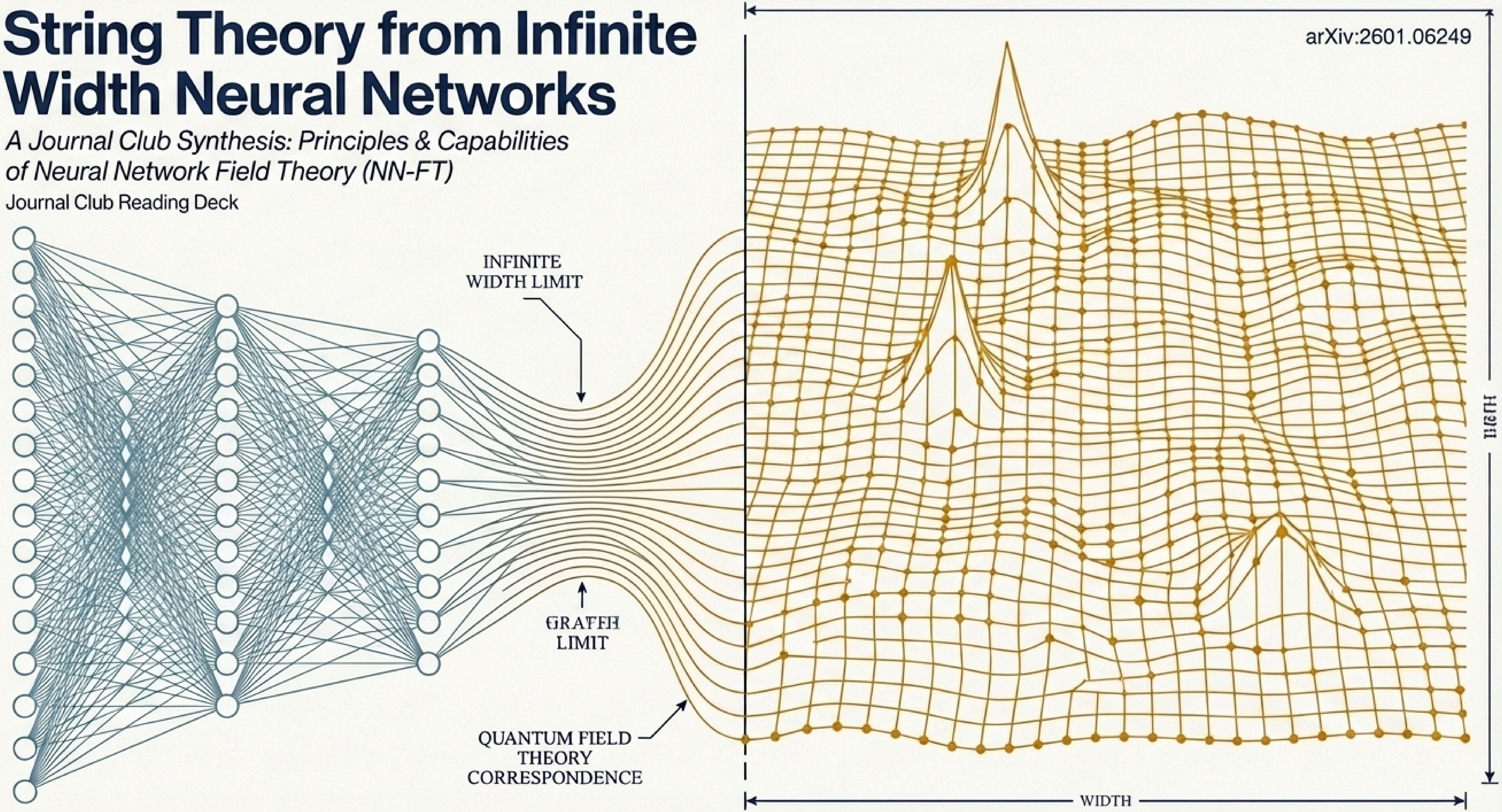
String Theory from Infinite Width Neural Networks

A Journal Club Synthesis: Principles & Capabilities of Neural Network Field Theory (NN-FT)

Journal Club Reading Deck

arXiv:2601.06249



Redefining the Field: From Path Integrals to Parameter Space

Traditional QFT

$$Z[J] = \int \overbrace{\mathcal{D}\phi}^{\text{Field Space}}\, \overbrace{e^{-\frac{1}{2}\phi K \phi + J\phi}}^{\text{Action } S[\phi]} = \exp\left(\frac{1}{2}\int J(x)G(x,y)J(y)\right)$$

$$\xrightarrow{\text{Equivalence}}$$

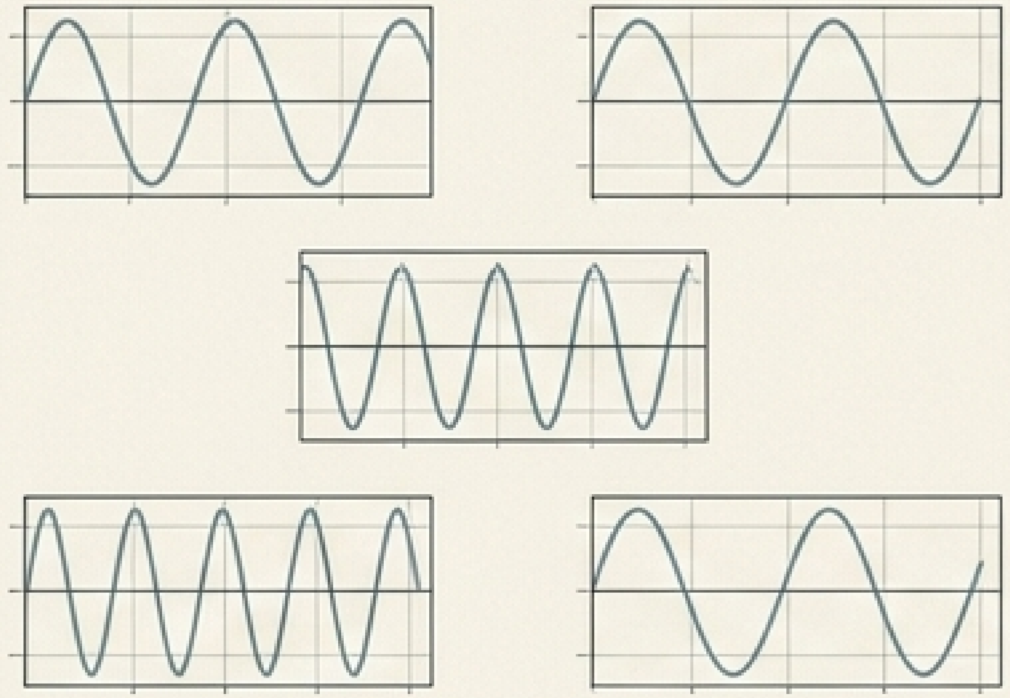
Neural Network Field Theory (NN-FT)

$$\underbrace{\mathbb{E}_{\theta}}_{\text{Parameter Space}}\left[\overbrace{\underbrace{f_{\theta}}_{\text{Architecture}}}^{\text{Architecture}}(x)\right]$$

A field theory is now defined entirely by two pieces of data:
1. An Architecture (Kinematics). **2.** A Parameter Density (Dynamics).

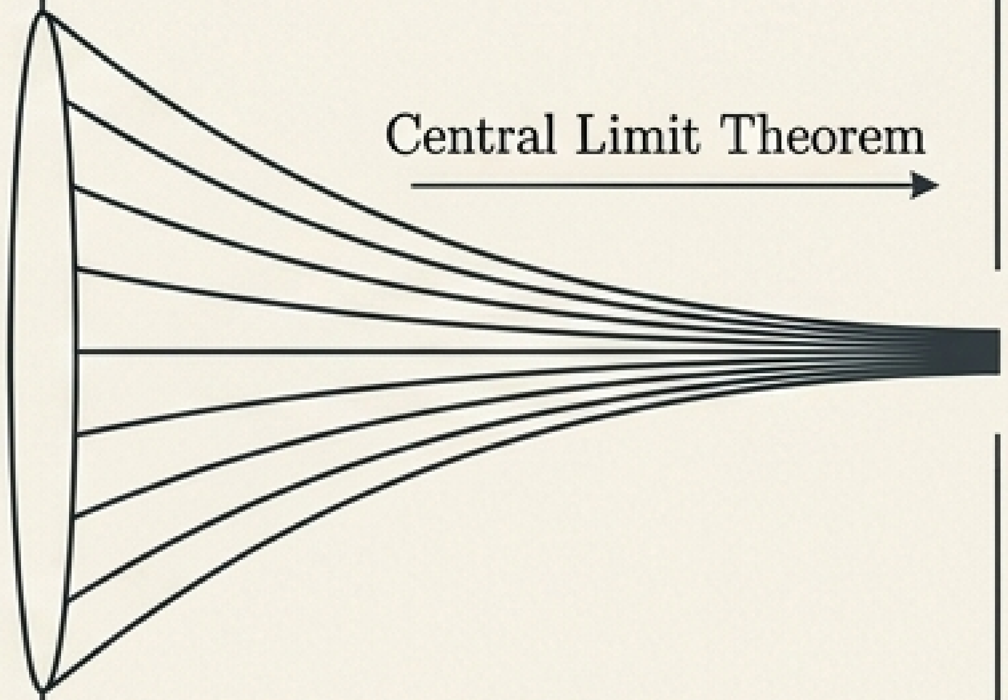
The Bridge to Free Fields: Infinite Width Yields Gaussian Processes

Finite Nodes

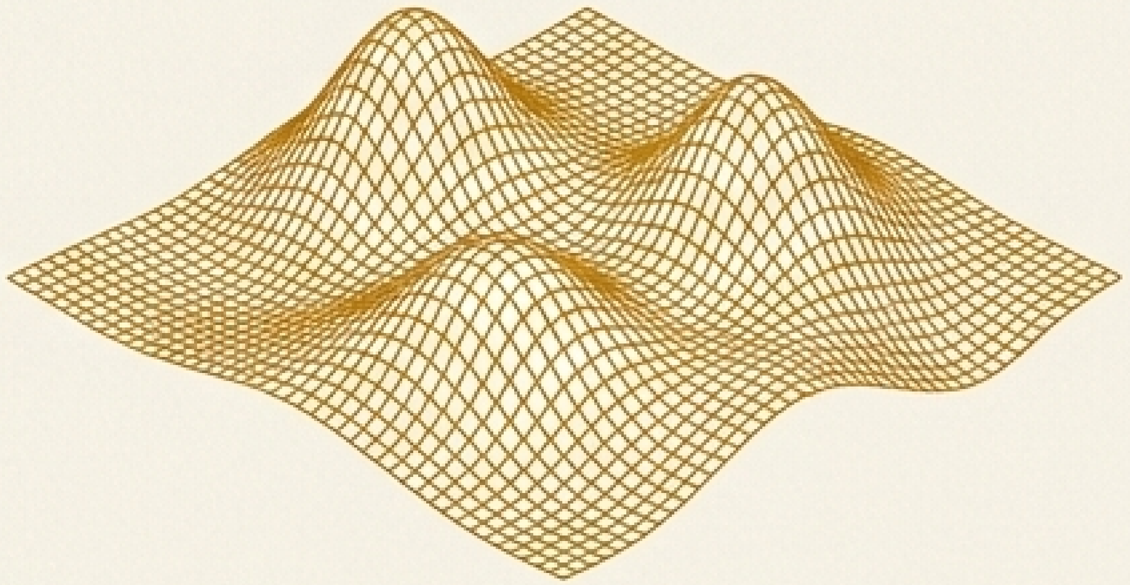


Single Neurons =
Independent random variables.

The $N \to \infty$ Limit



The Result



$$\phi(x) \sim GP(m(x), k(x, y))$$

**Infinite width exactly realizes a non-interacting (Free) Quantum Field Theory.
The network output distribution is strictly Gaussian.**

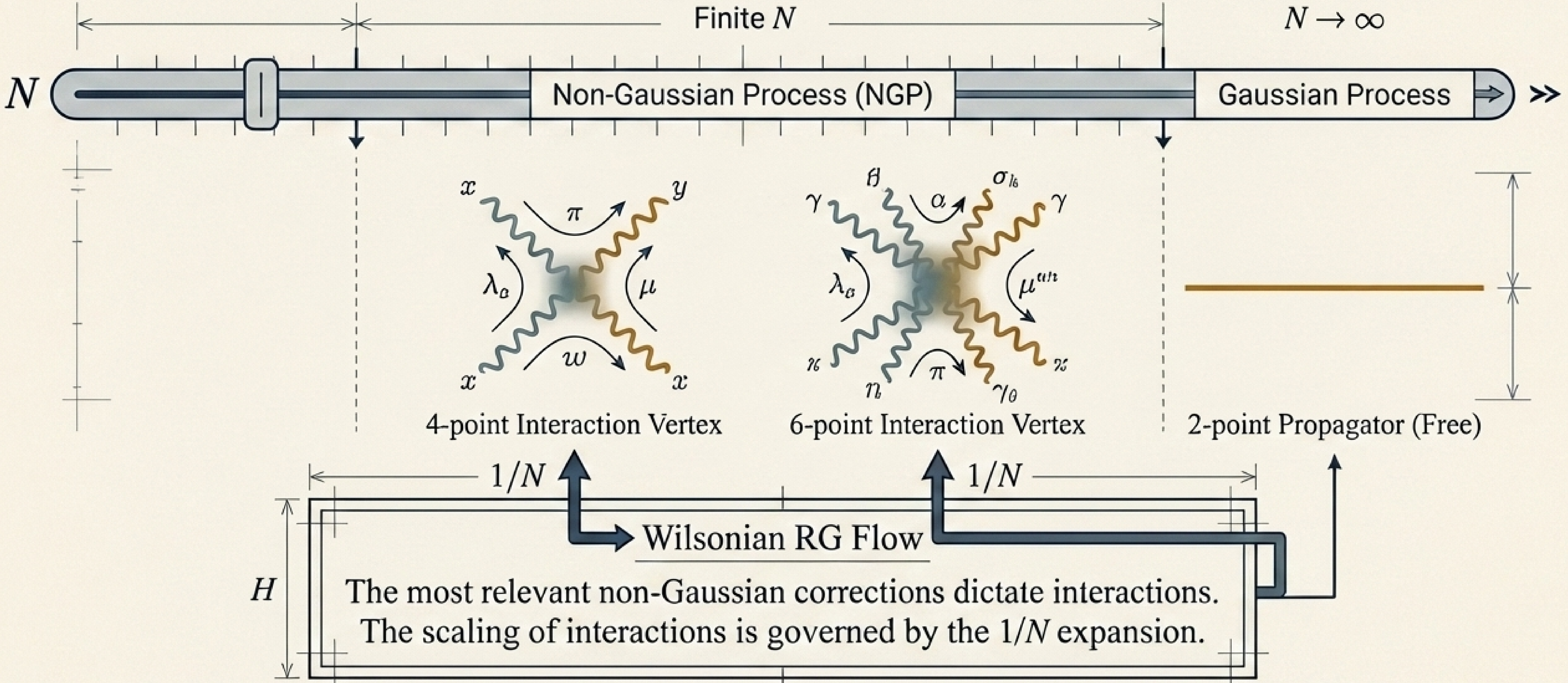
The NN-FT Rosetta Stone

Machine Learning (Cyan)		Quantum Field Theory (Amber)
Input x	$\longleftrightarrow$	Spacetime / Momentum point
Network Output $f(x)$ / $\phi(x)$	$\longleftrightarrow$	Interacting Field
Infinite Width ($N \to \infty$)	$\longleftrightarrow$	Free Theory (Gaussian Process)
Finite Width ($1/N$)	$\longleftrightarrow$	Interaction Strength (Non-Gaussianities)
Variance of Weights (σ_a^2)	$\longleftrightarrow$	Coupling Constant / String Tension (α')
Kernel $K(x, y)$	$\longleftrightarrow$	Feynman Propagator $\langle\phi(x)\phi(y)\rangle = k(x, y)$

This dictionary allows us to 'read' particle physics directly from neural network architectures.

Finite Width Generates Particle Interactions

arxiv : 2008.08601



Architecture as Geometry: Defining the Bosonic String

$$X^{\mu}(z) = \frac{C}{\sqrt{N}} \sum_{i=1}^{N} \frac{a_i^{\mu}}{|W_i|} \cos\left(\frac{1}{2}(W_i z + \bar{W}_i \bar{z}) + c_i\right)$$

Network Output = Spacetime Embedding

The output of the neural network represents the position of the string $X^{\mu}(z)$ in target spacetime. The complex coordinate z on the worldsheet is the Input.

Parameter Definitions: Weights & Symmetries

a_i^{μ}: Output Weights $\sim$ Isotropic Gaussian

Ensures Spacetime Rotational Invariance (Lorentz symmetry).

W_i: Frequencies $\sim \mathrm{Uniform}(B^2_{\Lambda} \setminus B^2_{\epsilon})$

Momentum modes on the annulus, defined over a fundamental domain in the complex plane.

c_i: Bias $\sim \mathrm{Uniform}[-\pi, \pi]$

Enforces Worldsheet Translation Invariance.

Activation = Mode Expansion & Translation Invariance

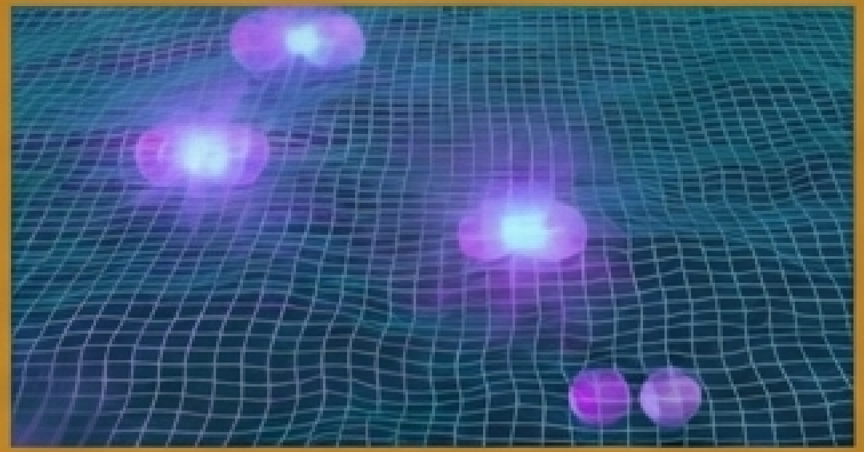
The cosine activation function represents the mode expansion of the string. The **uniform bias** $c_i \sim [-\pi, \pi]$ enforces **Translation Invariance** on the worldsheet, a key geometric symmetry.

The string is an ensemble of neural networks. The architecture itself enforces the geometric symmetries of the worldsheet.

Tuning the String: Variance Sets the Tension

$$\langle X^\mu(z)X^\nu(w)\rangle = \frac{\sigma_a^2 C^2}{2}\mathbb{E}_W\left[\frac{\cos[(W\Delta z + \bar{W}\Delta\bar{z})/2]}{|W|^2}\right]\delta^{\mu\nu},$$

$$\langle X^\mu X^\nu\rangle(r) = \alpha' \int_\epsilon^\Lambda d|W|\,\frac{J_0(r|W|)}{|W|}\delta^{\mu\nu} = \alpha'\delta^{\mu\nu}\left[-\log r + \log\frac{2\Lambda}{\epsilon e^\gamma} + \dots\right],$$



Physical String Tension

The string length $l_s \propto \sqrt{\alpha'}$ is exclusively determined by σ_a, the standard deviation of the neural network's output weights.

Virasoro-Shapiro Amplitudes Without Path Integrals

Vertex Operators

$$V(p) = g_s Z(p) \int d^2z \, e^{ip\cdot X(z)}$$

Insert wave/excited states $V(p)$

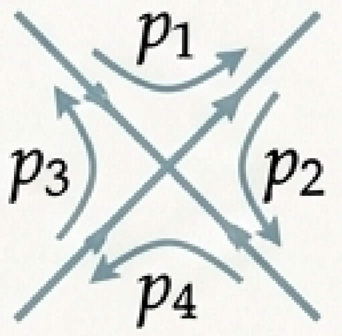
Parameter Averaging

$$\cancel{\int D\phi \, e^{iS[\phi]}} \Rightarrow \left\langle \prod V(p_i) \right\rangle$$

Replace Path Integral with Parameter Expectation

The Amplitude

$$A^{(4)} = \frac{g_s^2 \delta^D\left(\sum_r p_r\right)}{\mathrm{Vol}(\mathrm{SL}(2,\mathbb{C}))} \int \left(\prod_{\alpha=1}^{4} d^2 z_\alpha\right) \prod_{r<s} |z_r - z_s|^{\alpha' p_r \cdot p_s},$$



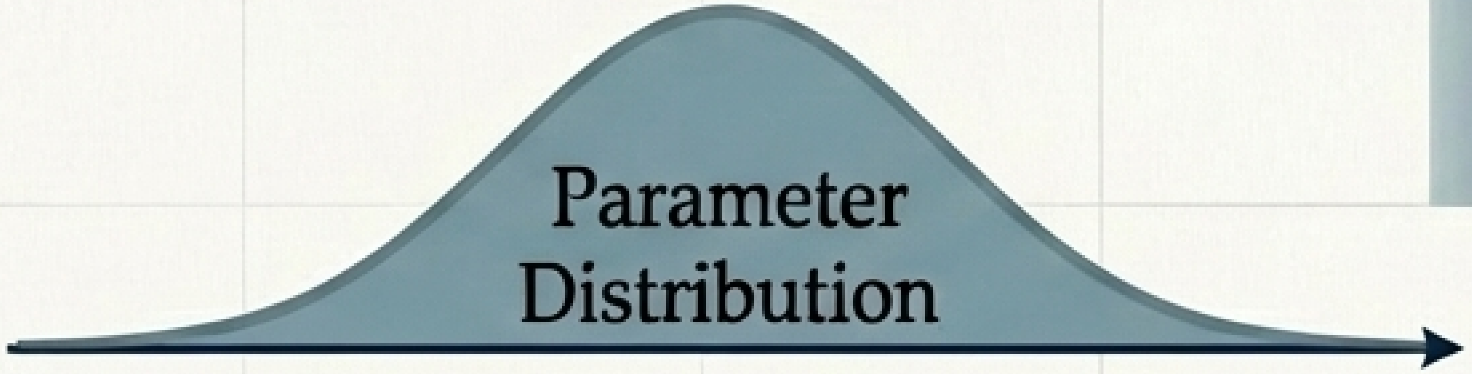
Crossing symmetry, Regge behavior, and UV softness emerge natively from the neural network's parameter statistics.

arxiv:2512.24420

Enforcing Local Symmetry: The Log-Kernel Architecture

Standard Architectures

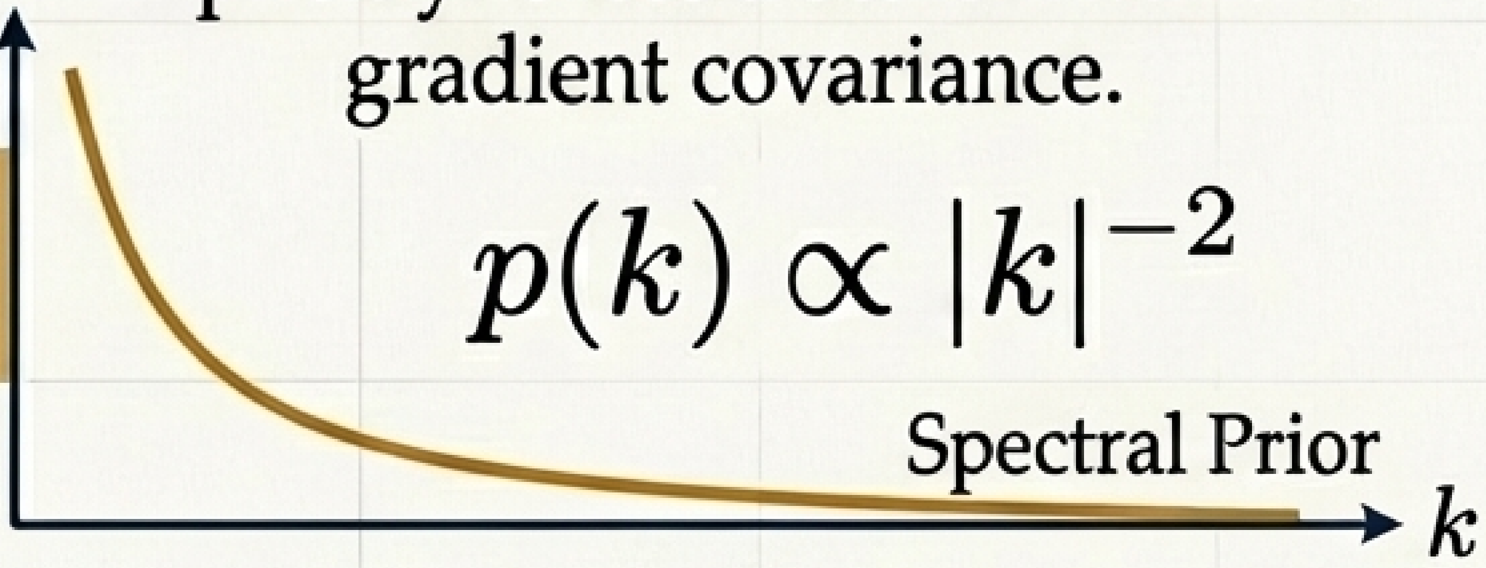
Fails to produce a local stress-energy tensor. Global scale invariance only.



Engineering the Prior

Explicitly forces a scale-invariant gradient covariance.

$$p(k) \propto |k|^{-2}$$



Theoretical Central Charge: $c = 1$

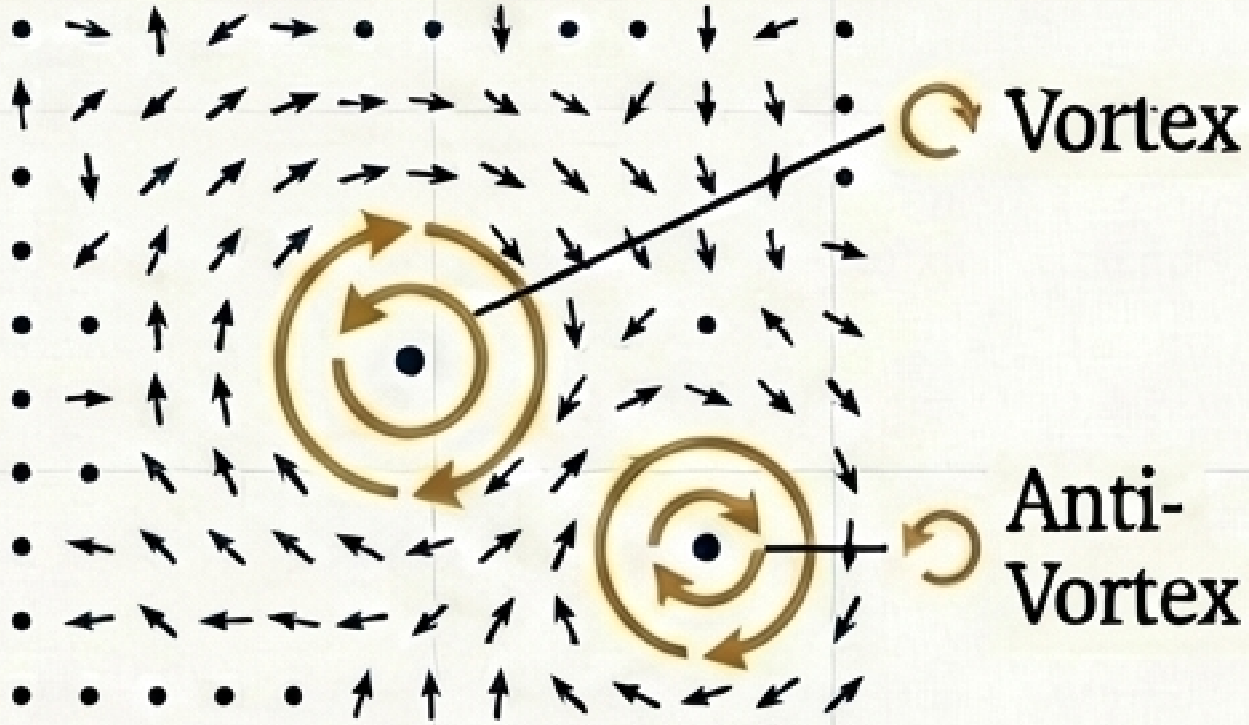
Experimental Central Charge: $c_{\text{exp}} = 0.9958 \pm 0.0196$

By carefully engineering the probability measure of the network ensemble itself, we can force the hidden layer to act as a conformal reservoir.

arxiv:2604.02313

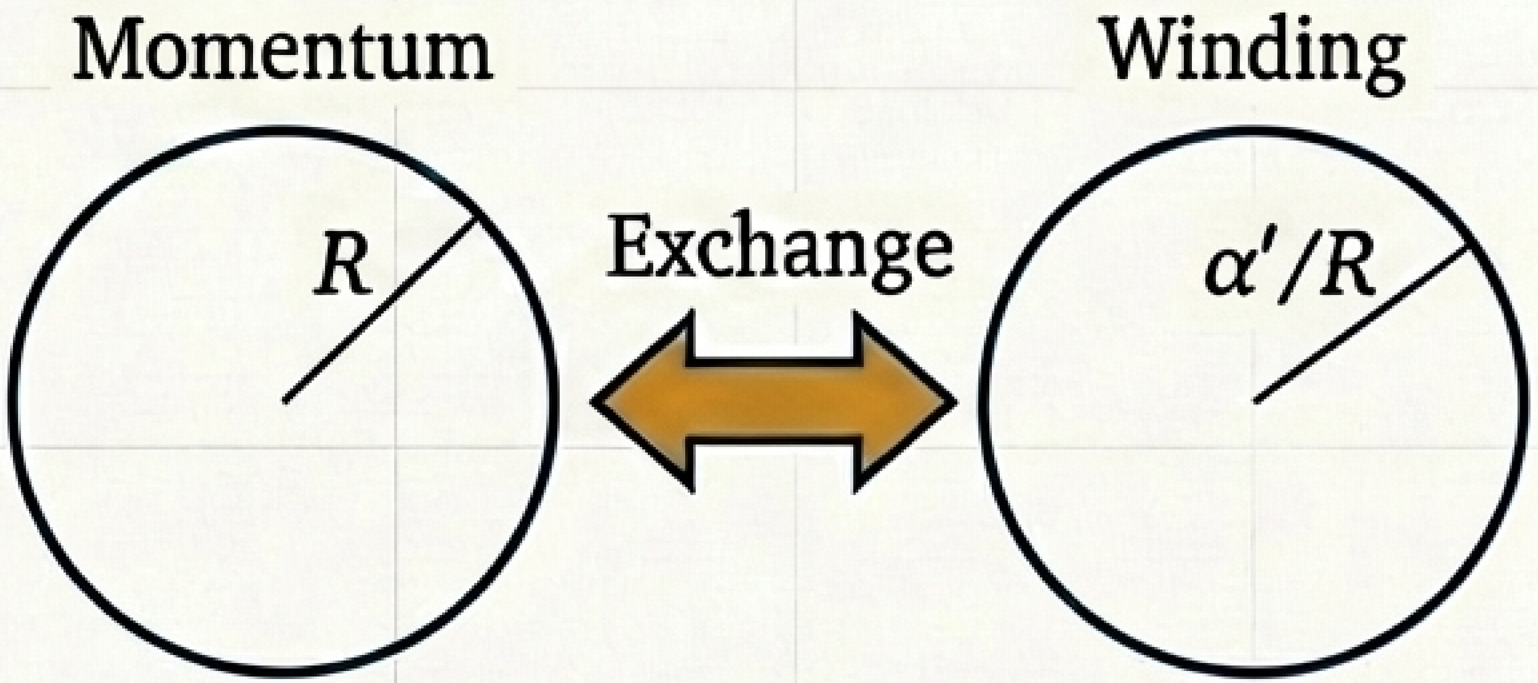
Capturing Global Structure: Topology & Dualities

A BKT Transition



Incorporating discrete network parameters unlocks the Berezinskii-Kosterlitz-Thouless transition at high temperatures.

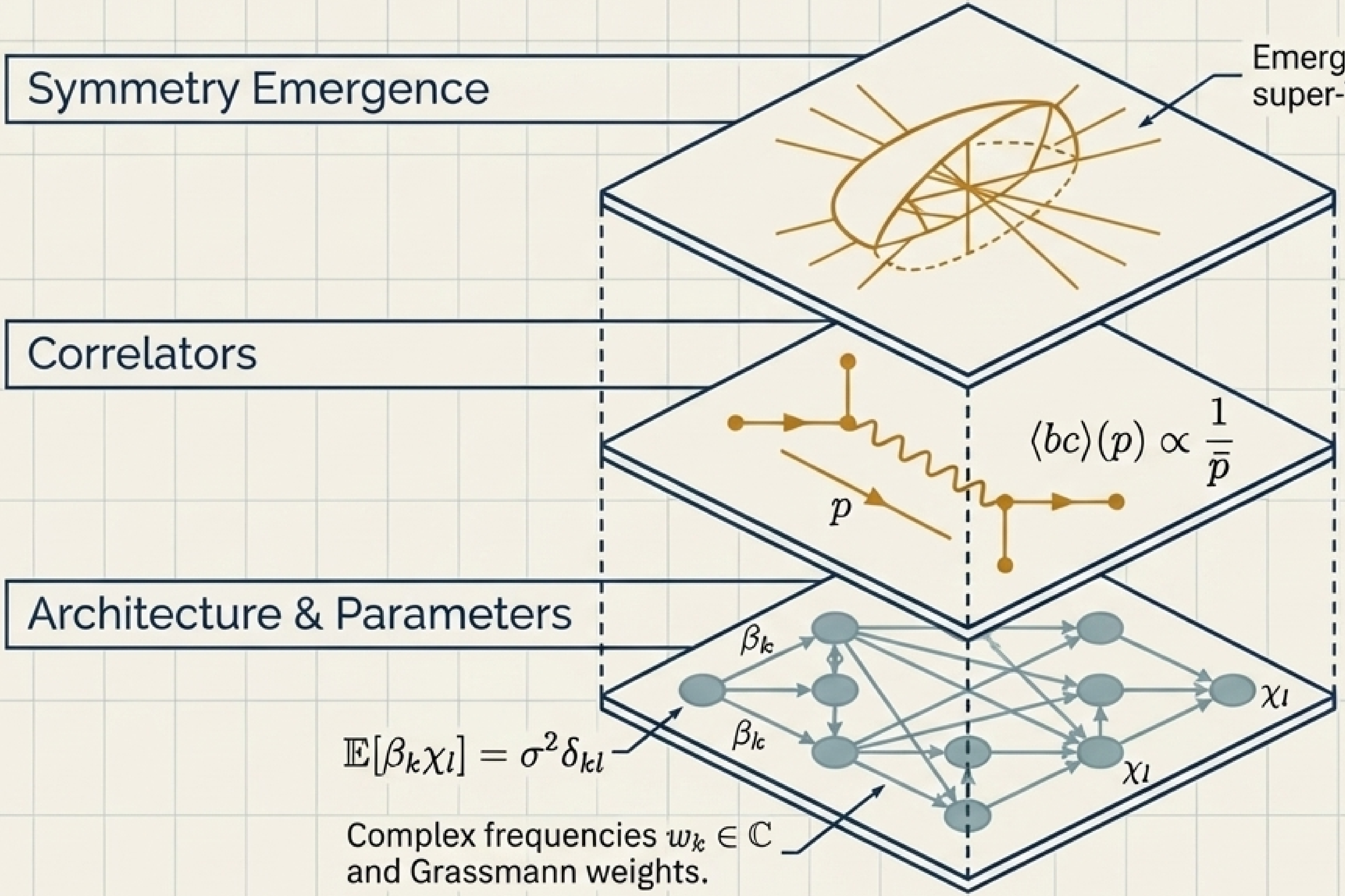
B T-Duality



The network recovers the Buscher rules and exhibits exact invariance under the exchange of momentum and winding.

Continuous neural sectors encode local degrees of freedom; discrete latent variables encode topological and flux sectors.

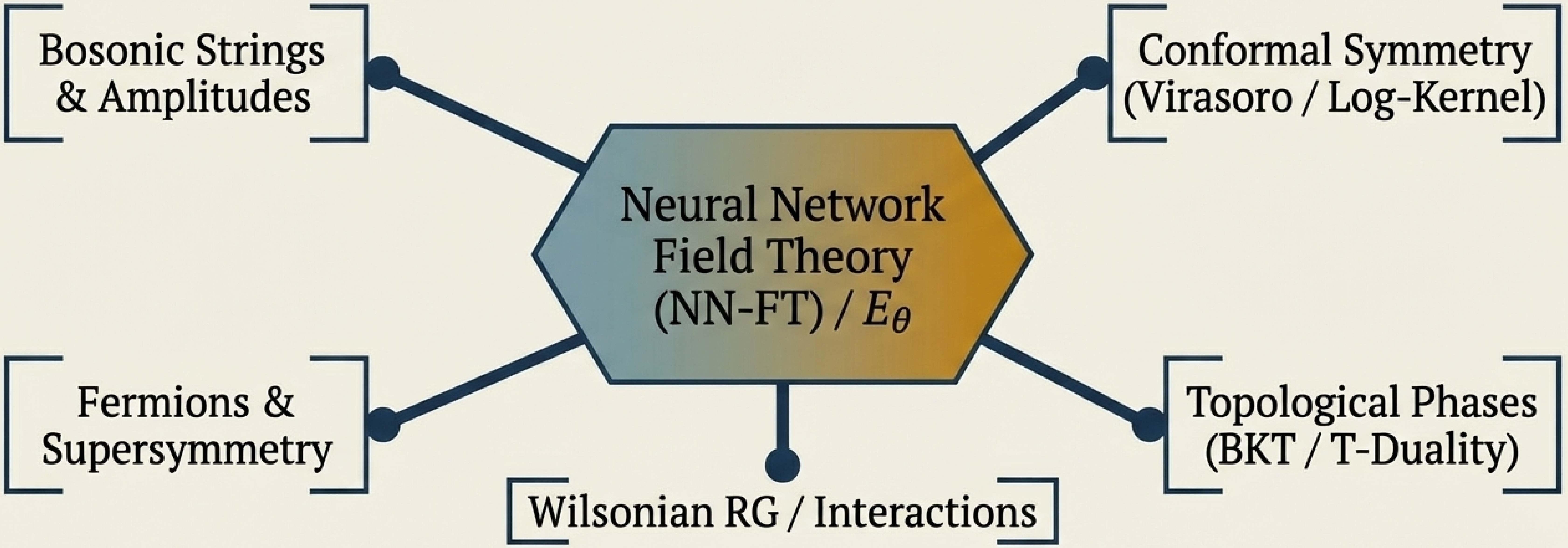
Expanding the Particle Zoo: Fermions & Ghosts



NN-FT is not limited to commuting scalars. Grassmann-valued neural networks perfectly replicate fermionic statistics and supercurrent correlators.

fermionc statistics and supercurrent correlators threann ooned neural networks, experimenti theory.

Synthesis: The NN-FT Ecosystem

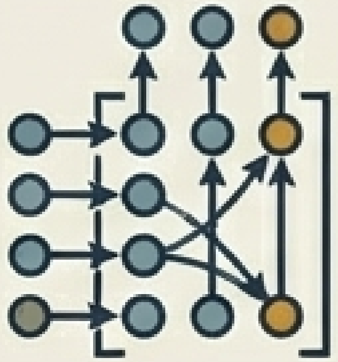


Any architecture with a GP limit—MLPs, CNNs, or Transformers—can theoretically be mapped to a corresponding field theory. The architecture dictates the kinematics; the prior dictates the dynamics.

Future Horizons: Where Do We Go From Here?

Architecture Exploration

What novel scattering amplitudes emerge if we replace MLPs with Transformers or Transformers or CNNs? What is the physics of self-attention?



Exotic Dualities

If T-duality is natively encoded, can Mirror Symmetry or S-duality be S-duality be programmed into network priors?

Scientific ML

Can we use scale-invariant architectures (like the Log-Kernel) as inductive biases to drastically improve training on critical phenomena or turbulence data?

"The worldsheet path integral is now an integral over neural network parameter space."