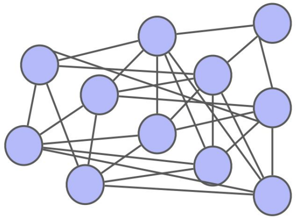
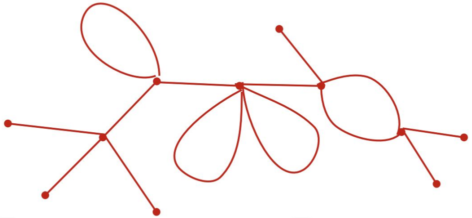
# A Study of Neural Network Field Theories


## Anindita Maiti



King's College London | TPPC Seminar
Oct 12, 2022

Based on arxiv: 2008.08601, 2106.00694 & 22xx.xxxxx (w/ Jim Halverson, Matt Schwartz, Mehmet Demirtas & Keegan Stoner)

# What are Neural Networks?

Backbones of Deep Learning.

Outputs are functions of inputs, with continuous learnable parameters $\boldsymbol{\theta}$ and discrete hyperparameter N.

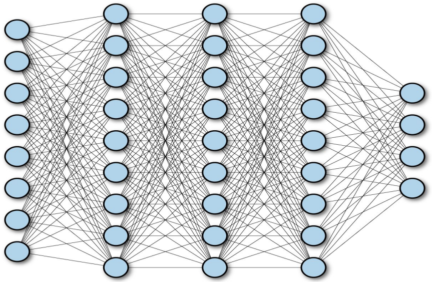
$$f_{\theta,N} : \mathbb{R}^{d_{\text{in}}} \to \mathbb{R}^{d_{out}}$$

Generate NN outputs multiple times, outputs get drawn from same distribution.

## Fully Connected Feedforward NN :

→ “depth”

→ Width N



input → hidden → …
→ hidden→ output

$$z_i^l(x) = b_i^l + \sum_{j=1}^{N} W_{ij}^l x_j^l(x)$$

$$x_j^l = \sigma(z_j^{l-1}(x))$$

Statistical perspective: Field Theories are defined by distributions on field / function space (via Feynman path integral).

Action S[ϕ] is the ‘log-likelihood’

$$Z = \int D\phi\, e^{-S[\phi]}$$

# Punchlines

- Ensembles of Neural Network outputs behave as Euclidean Field Theories.
- Neural Network Gaussian Processes (NNGP) correspond to Free Field Theories. Deviations from NNGP turn on interaction terms.
- Small & large deviations leads to weakly coupled & non-perturbative Neural Network Field Theories, respectively.
- NNs also have a dual “parameter distributions + architecture” framework.
- Symmetries, connected correlators, partition function, and locality via cluster decomposition can be studied in this dual framework; knowledge of action isn’t necessary.

# Motivations

- Ongoing work to find NN architectures, that would result in known Euclidean Field Theories, at initialization.

- Bypassing NN training / learning to generate desired Field Theories at initialization.

- If Osterwalder-Schrader axioms are satisfied by Euclidean correlators of NN output ensembles, then these NN architectures correspond to QFTs. [Halverson 2021]

- Combinations of such NN architectures would result in field theories with actions similar to that of the Standard Model.

# References & Related Works

## Based on:

1. arXiv:2008.08601
2. arXiv:2106.00694
3. arXiv:22xx.xxxxx (to appear soon)

Jim Halverson

Matt Schwartz

## Related Works:

[Halverson 2021],
[Erbin, Lahoche, Dine 2021],
[Grosvenor, Jefferson 2021],
[Lee, Bahri, Novak, Schoenholz, Pennington, Sohl-Dickstein 2017],
[Yang 2019],
[Roberts, Yaida, Hanin 2021],
[Yaida 2019].

Mehmet Demirtas

**Keegan Stoner**

# Outline

- Free Field Theories in Neural Networks
- Wilsonian Effective Field Theories in Neural Networks
- Non-Perturbative Neural Network Field Theories
- Study of NN Field Theories via Dual Parameter Space

# Free Field Theories in Neural Networks

# Free Field Theories in Neural Networks

## Limit N $\longrightarrow \infty$ and Independent Parameters: NN output ensembles become a Gaussian Process (NNGP), by Central Limit Theorem (CLT).*

| GP / asymptotic NN | Free QFT |
|---|---|
| input $x$ | external space or momentum space point |
| kernel $K(x_1, x_2)$ | Feynman propagator |
| asymptotic NN $f(x)$ | free field |
| log-likelihood | free action $S_{\mathrm{GP}}$ |

## Neural Network Gaussian Process:

PDF of NN output ensemble:

$$P[f] \sim \exp\left[-\frac{1}{2}\int d^{d_{\mathrm{in}}}x\, d^{d_{\mathrm{in}}}x'\, f(x)\Xi(x,x')f(x')\right]$$

$$\int d^{d_{\mathrm{in}}}x'\, K(x,x')\,\Xi(x',x'') = \delta^{(d_{\mathrm{in}})}(x-x'')$$

Resembles the PDF of free field theory, from Feynman path integral formalism.

## Define partition function:

$$Z = \int D\phi\, e^{-S[\phi]}$$

## Predict correlators:

$$G^{(n)}(x_1,\ldots,x_n) = \frac{\int df\;\, f(x_1)\ldots f(x_n)\, e^{-S}}{Z}$$

* NNGP references: [Neal], [Williams] 1990's , [Lee et al., 2017], [Matthews et al., 2018] , [Yang, 2019], [Yang, 2020]

# Wilsonian EFT in Neural Networks

# Wilsonian EFT in Neural Networks

| NGP / finite NN | Interacting QFT |
|---|---|
| input $x$ | external space or momentum space point |
| kernel $K(x_1, x_2)$ | free or exact propagator |
| network output $f(x)$ | interacting field |
| non-Gaussianities | interactions |
| non-Gaussian coefficients | coupling strengths |
| log probability | effective action $S$ |

Close to the NNGP, Wilsonian EFT* describes NN output ensembles.

$$S = S_{\mathrm{GP}} + \Delta S$$

$$\Delta S = \int d^{d_{\mathrm{in}}}x \, \left[ g\, f(x)^3 + \lambda\, f(x)^4 + \alpha\, f(x)^5 + \kappa\, f(x)^6 + \dots \right]$$

Following EFT, some couplings are more relevant than others. Further, relative 1/N scaling between couplings.

Closeness to NNGP $\longrightarrow$ NN parameter framework becomes complex, due to a large number of parameters.

But EFT action gets fewer relevant interaction terms.

$$G^{(n)}(x_1, \dots, x_n) = \frac{\int df \;\, f(x_1) \dots f(x_n)\, e^{-S}}{Z_0}$$

Predict correlators. Run NN expt to determine couplings and interaction terms. $\longleftarrow$ NN phenomenology.

* [Erbin, Lahoche, Dine 2021] gives an alternate approach for NN ensembles close to the GP.

# Non-Perturbative Neural Network Field Theories

# Non-Perturbative Neural Network Field Theories

Small width and/or large parameter correlations violate Central Limit Theorem by large amounts.

Leads to non-perturbative NN field theories, with the action often unknown.

$$f_{\theta,N}(x) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} h_i(x)$$

$h_i(x)$: output of i-th neuron in final hidden layer.

NN output ensembles can be studied using NN architecture and parameter distributions, too.

**Field Space**

$$G^{(n)}(x_1, \ldots, x_n) = \frac{\int df \; f(x_1) \ldots f(x_n) \, e^{-S}}{Z_0}$$

**Parameter Space**

$$\mathbb{E}[f(x_1) \cdots f(x_n)]$$
$$= \frac{1}{N^{h/2}} \sum_{i_1, \cdots, i_n = 1}^{N} \int DhP(h) h_{i_1}(x_1) \cdots h_{i_n}(x_n)$$

Properties and observables, e.g. symmetries, partition functions etc. can be studied without knowledge of actions.

# Study of NN Field Theories via Dual Parameter Space

# Symmetry via Duality

NNGP: symmetries of 2-pt function determine symmetries of NN distribution.

E.g. Mean-free SO($d_{out}$) invariant parameter distributions lead to SO($d_{out}$) invariant free field theory action in NN.

$$G^{(2)}_{i_1 i_2}(x_1, x_2) = \delta_{i_1 i_2} K(x_1, x_2)$$

R ∈ SO($d_{out}$), output transforms as $f_i \mapsto R_{ij} f_j$

$$G^{(2n)}_{i_1,\dots,i_{2n}}(x_1, \dots, x_{2n})$$
$$= \sum_{P \in \mathrm{Wick}(2n)} \delta_{i_{a_1} i_{b_1}} \dots \delta_{i_{a_n} i_{b_n}} K(x_{a_1}, x_{b_1}) \dots K(x_{a_n}, x_{b_n})$$

Break Central Limit Theorem assumptions: n>2 correlators receive EFT corrections.

NN action unknown: symmetries can't be deduced in field space.

Study NN correlators using NN architecture and parameter distributions.

NN action remains invariant

$$D[\Phi f]\, e^{-S[\Phi f]} = Df\, e^{-S[f]}$$

if transformations $f'(x) = \Phi(f(x'))$ leave correlators invariant.

# Symmetry via Duality

$$\mathbb{E}[f(x_1)\dots f(x_n)] = \frac{1}{Z_f}\int Df\, e^{-S[f]}\, f(x_1)\dots f(x_n)$$

$$= \frac{1}{Z_f}\int D[\Phi f]\, e^{-S[\Phi f]}\, \Phi(f(x_1'))\dots \Phi(f(x_n'))$$

$$= \frac{1}{Z_f}\int Df\, e^{-S[f]}\, \Phi(f(x_1'))\dots \Phi(f(x_n')) = \mathbb{E}[\Phi(f(x_1'))\dots \Phi(f(x_n'))]$$

Absorb transformations of correlators into transformations of parameters.

Invariance of parameter distributions leads to invariance of NN action S[f].

Symmetries of NN input and output layers $\longrightarrow$ spatial symmetries and internal symmetries of fields, respectively.

**Examples:**

(a) **SO($d_{out}$) Output Symmetry:** Final linear layer parameters drawn from mean-free SO($d_{out}$) invariant distributions.

$$f_i(x) = W_{ij}g_j(x) + b_i \qquad f_i \mapsto R_{ij}f_j$$

$$P_W = P_{R^{-1}\tilde{W}} = P_{\tilde{W}} \qquad P_b = P_{R^{-1}\tilde{b}} = P_{\tilde{b}}$$

$$R \in SO(d_{\text{out}})$$

**(b) SO($d_{in}$) Input Symmetry:** First linear layer parameters drawn from mean-free SO($d_{in}$) invariant distributions.

$$f_i(x) = g_{ij}(W_{jk}x_k)$$

$$R \in SO(d_{\text{in}}) \qquad x_i \mapsto x_i' = R_{ij}x_j$$

# Connected Correlators via Duality

**Cumulant Generating Functional (CGF)** of field theories in NNs, in terms of cumulants / connected correlators expressed in NN architecture framework.

NN output as a field or as a sum over neuron contributions.

$$f_{\theta,N}(x) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} h_i(x)$$

Correlated parameter distributions.

$$P(h|\vec{\alpha}) \neq \prod_i P_i(h_i) \qquad \vec{\alpha} = \{\alpha_1, \cdots, \alpha_q\}$$

CGF of NN field theory:

$$W_f[J] = \sum_{r=1}^{\infty} \int \prod_{i=1}^{r} d^{d_{\text{in}}} x_i \frac{J(x_1)\dots J(x_r)}{r!} G^{(r)}_{\text{con},f}(x_1,\dots,x_r)$$
$$= \log\left[\int Dh\, P(h|\vec{\alpha})\, e^{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int dx\, h_i(x) J(x)}\right]$$

Finite width, no parameter correlations:

$$G^{(r)}_{\text{con},f}(x_1,\cdots,x_r) = \frac{G^{(r)}_{\text{con},h_i}(x_1,\cdots,x_r)}{N^{r/2-1}}$$

**Finite width, parameters correlated:** each cumulant in NN field theory receives contributions from cumulants of all neurons.

# Partition Function via Duality

**Edgeworth Expansion:** Inverse Fourier transform of the CGF to obtain NN field theory PDF, and then the field theory partition function.

$$W_f[J] = \log Z_f[J] \qquad Z_f[J] = \int Df P_f \, e^{i\int d^{d_{\rm in}}x\, J(x) f(x)}$$

PDF of Non-Perturbative NNFT ⟶ perturbative expansions around PDF of Free NNFT.

Coefficients of terms in the series expansion of PDF are given by cumulants, with relative 1/N scalings.

$N \to \infty, \vec{\alpha} \to \vec{0}$ : Free NN field theory PDF.

$$P_f = \exp\Big(-\frac{1}{2}\int d^{d_{\rm in}}x_1 d^{d_{\rm in}}x_2\, f(x_1)\Xi(x_1,x_2)f(x_2)\Big)$$

$$\int d^{d_{\rm in}}x'\, \Xi(x_1,x')G^{(2)}_{\rm con,f}(x',x_2) = \delta^{d_{\rm in}}(x_1-x_2)$$

$$P_f = \int DJ \exp\Big(\sum_{r=3}^{\infty}\frac{(-i)^r}{r!}\int\prod_{i=1}^{r} d^{d_{\rm in}}x_i G^{(r)}_{\rm con,f}(x_1,\cdots,x_r)\frac{\partial}{\partial f(x_1)}\cdots\frac{\partial}{\partial f(x_r)}\Big)$$
$$\times \exp\Big(-i\int d^{d_{\rm in}}x\, J(x)f(x) - \frac{1}{2}\int d^{d_{\rm in}}x_1 d^{d_{\rm in}}x_2\, J(x_1)G^{(2)}_{\rm con,f}(x_1,x_2)J(x_2)\Big)$$

**NN Field Theory partition function is obtained from $P_f$**

# Locality via Cluster Decomposition

We can study locality via cluster decomposition, even in absence of action, using cumulants / connected correlators in NN parameter space.

$$\sum_{r=1}^{\infty} \int \prod_{i=1}^{r} d^{d_{\text{in}}} x_i \frac{J(x_1)\dots J(x_r)}{r!} G^{(r)}_{\text{con},f}(x_1,\dots,x_r) = \log\left[\int \left(\prod_{i=1}^{N} Dh_i\right) P(h|\vec{\alpha}) e^{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int d^{d_{\text{in}}}x\, h_i(x) J(x)}\right]$$

Take any two subsets of input points ($x_1$, …, $x_r$) infinitely apart.

If all connected correlators (r>1) decay to 0 exponentially or faster, then the NN field theory is local.

**Choose NN architecture and parameter distributions such that all cumulants, for r>1, are local, by cluster decomposition method.**

# Conclusions

- NN output ensembles have dual descriptions by field space and NN parameter space.
- Free Field theories, Wilsonian EFTs and Non-Perturbative Non-Lagrangian Field Theories arise in NNs at GP limit, small, and large deviations from GP limit, respectively.
- When Euclidean NN output correlators satisfy Osterwalder-Schrader axioms, NNs define QFTs. [Halverson 2021]
- More parameters in NN towards GP limit, fewer relevant interaction terms in NN Field Theory action.
- Small width and / or large parameter correlations lead to non-perturbative NN field theories with unknown actions.
- Symmetry, connected correlators, partition functions, and locality via cluster decomposition can be studied without necessitating the field action, using dual NN parameter framework.

# Thank You!

Questions?

maiti.a@northeastern.edu

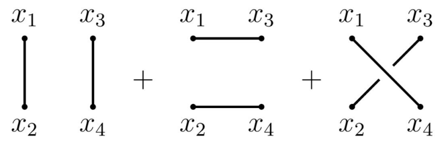
# Extra Slides

# Free Field Theories in Neural Networks

## Introduce Feynman diagrams for NN correlators

$$\Delta G^{(n)}(x_1,\dots,x_n) = G^{(n)}(x_1,\dots x_n) - G^{(n)}_{\mathrm{GP}}(x_1,\dots,x_n)$$

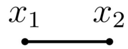
$$\Delta G^{(4)} = \frac{1}{n_{\mathrm{nets}}}\sum_{\alpha}^{n_{\mathrm{nets}}} f_\alpha(x_1)f_\alpha(x_2)f_\alpha(x_3)f_\alpha(x_4) - \Big[\quad\Big]$$




$$m_n = \Delta G^{(n)}/G^{(n)}_{\mathrm{GP}}$$

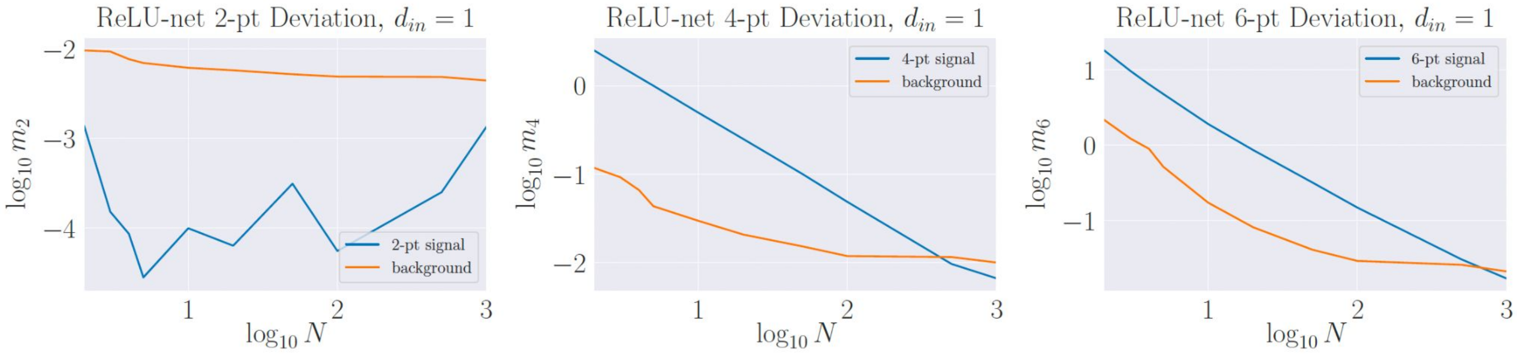
$$G^{(2)}_{\mathrm{GP}}(x_1,x_2) = K(x_1,x_2)$$

$$=$$




Test at various N; 100 expt, each with $10^5$ Nets.

ReLU-net: $\sigma(z) = \begin{cases} 0 & z<0 \\ z & z\geq 0 \end{cases}$




For n>2 ,

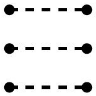
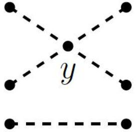
$$\Delta G^{(n)} \propto N^{-1}$$

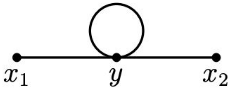
# Wilsonian EFT in Neural Networks

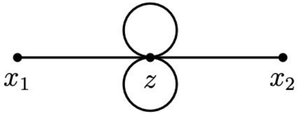
## Estimate $\lambda$ from 4-pt function expts

Fully connected feedforward NNs have an exact 2-pt function at all interaction strengths.

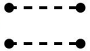
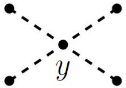
## Use $\lambda$ to predict 6-pt function

$G^{(6)}(x_1, x_2, x_3, x_4, x_5, x_6) = 15$  $- 360\,\lambda$ 


$G^{(2)}(x_1, x_2) \;=$  $- \lambda \Big[ 12$ 

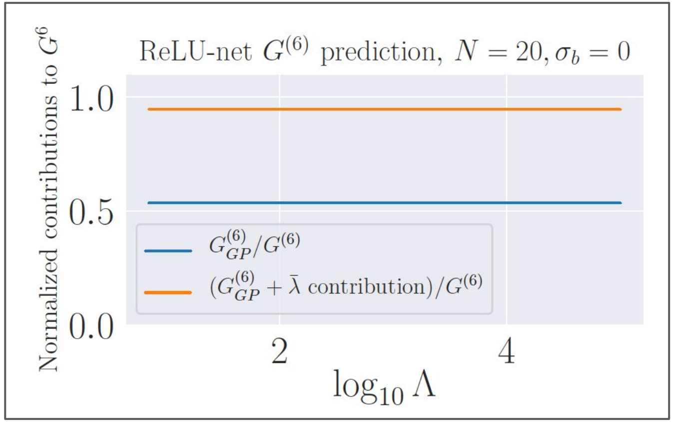
$\Big] - \kappa \Big[ 90$ 

$\Big] \;=$  $= \; K(x_1, x_2)$

$G^{(4)}(x_1, x_2, x_3, x_4) = 3$  $- 24\,\lambda$ 


$$= K(x_1, x_2)K(x_3, x_4) + K(x_1, x_3)K(x_2, x_4) + K(x_1, x_4)K(x_2, x_3)$$

$$- 24 \int_{-\Lambda}^{\Lambda} d^{d_{\text{in}}}y\,\lambda\,K(x_1, y)K(x_2, y)K(x_3, y)K(x_4, y)$$



# Wilsonian EFT in Neural Networks

Experimental correlation functions of NNs are invariant of choice of cutoff scales.

Cutoff dependence shows up as RG flow of couplings.

$$\frac{dG^{(n)}(x_1, \cdots, x_n)}{d\Lambda} = 0.$$

RG equation of quartic coupling in ReLU network: $\beta(\lambda) = \dfrac{\partial\lambda}{\partial \log\Lambda} = -(d_{\text{in}} + 4)\lambda$