Scalar Field Theories via Neural Networks at Initialization

Anindita Maiti

Email: amaiti@perimeterinstitute.ca

ML4Jets **Nov-9-2023**

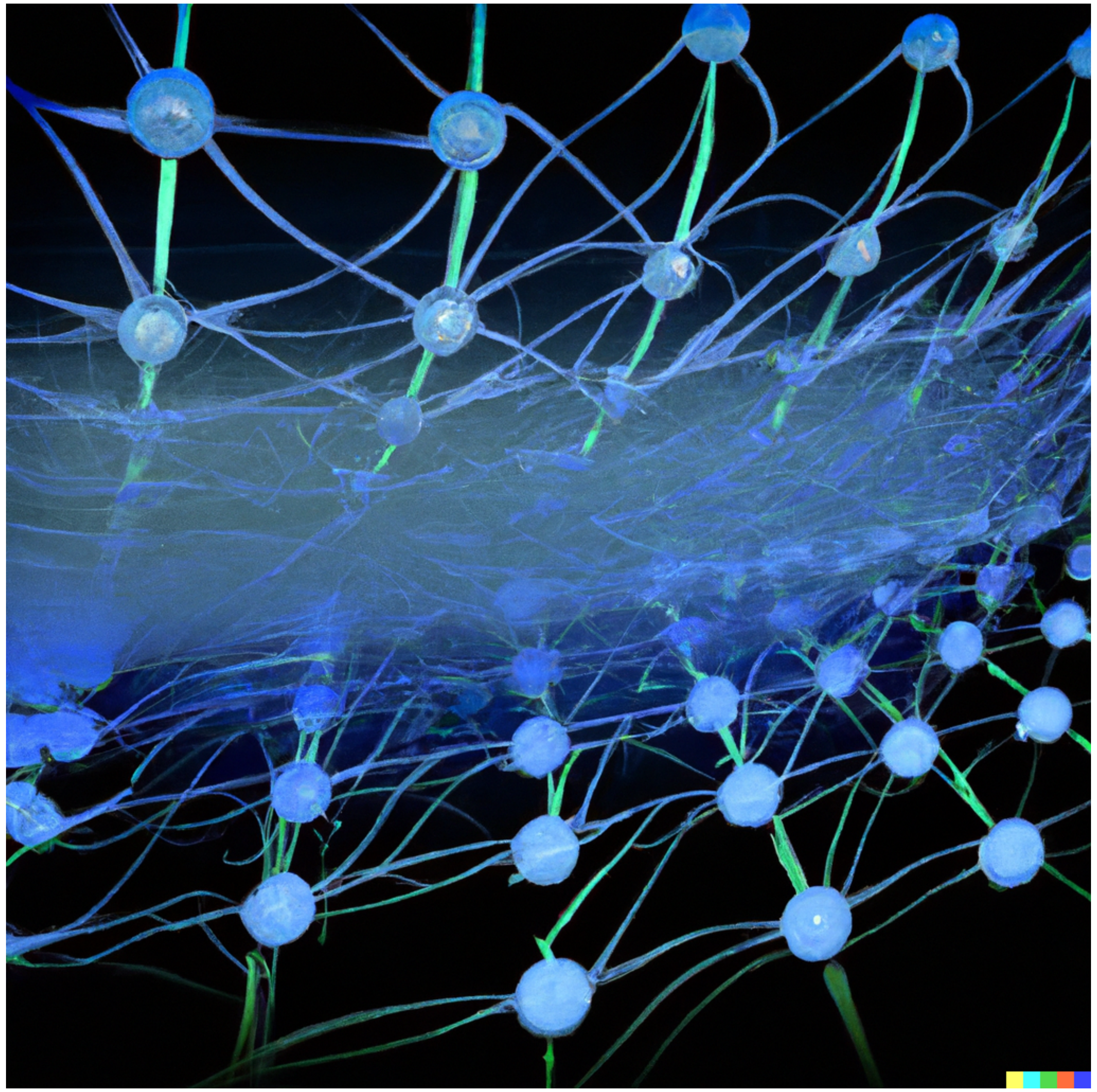
PERIMETER INSTITUTE

Based on [2307.03223], [2106.00694], [2008.08601] w/ Halverson, Schwartz, Demirtas, Stoner

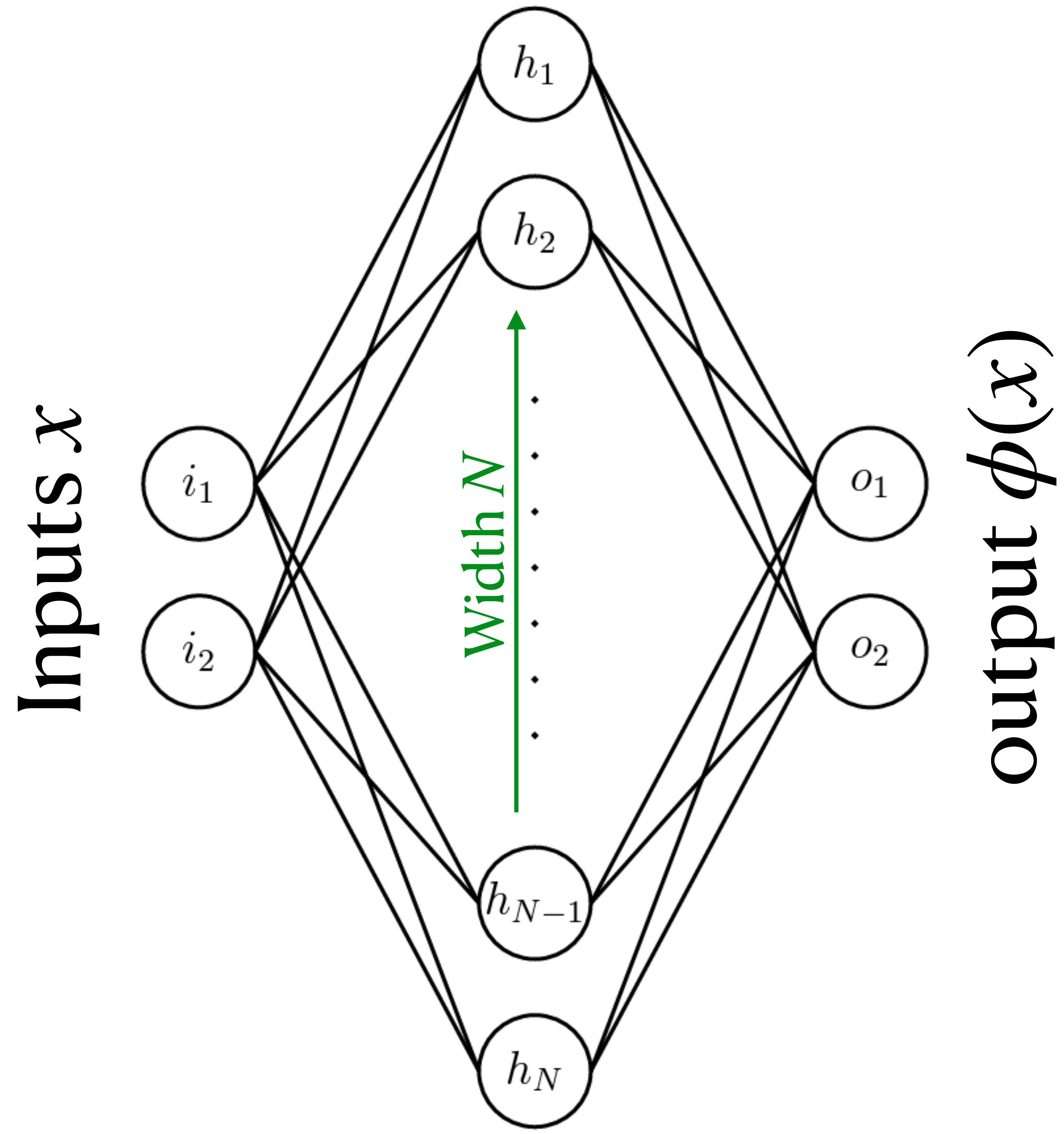
Overview & Motivation

- Deep Learning is widely used as a toolbox, to approximate systems.

- Can we come up with Neural Network architectures that can exactly represent some quantum field ϕ, without training? I.e. ab initio AI for qft?



Neural Network Field Theory Correspondence



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

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

- Initialize NN many times, do not train.
- Statistical distribution over NN outputs $\phi(x)$ can be cast in path integral formalism, as a field distribution.

Each initialized architecture corresponds to a unique field distribution!

Free vs. Interacting NN Field Distributions

Correlated parameter distributions

[Yaida 2019], [Yang 2019], [Dyer et al. 2019], [Naveh et al. 2021], [Roberts et al. 2022], [Hanin 2022], [Zavatone-Veth et al. 2022], [Pehlevan et al. 2021]

Related works!

- ML for field theories on lattice
- Correspondences between learning dynamics and RG and Hamiltonian evolution
- Correspondences between initialized NNs and field theory
- Theory of NN and deep learning at finite width, and other aspects of CLT breaking
- Theory of NN at asymptotically large width

[Halverson, A. M., Stoner 2020], [A.M., Halverson, Stoner 2021], [Jefferson et al. 2021], [Halverson 2022], [Erbin et al. 2022], [Banta et al. 2023], **[Halverson, Demirtas, A.M., Schwartz, Stoner 2023]**

[Albergo et al. 2019], [Hackett et al. 2021], [Bachtis et al. 2021], [de Haan et al. 2021], [Bachtis et al. 2022], [Abbott et al. 2022], [Gerdes et al. 2022]

[Berman et al. 2022], [Krippendorf et al. 2022], [De Luca et al. 2022], [Cotler et al. 2023]

Dissimilar NN parameter distributions

[Neal 1995]

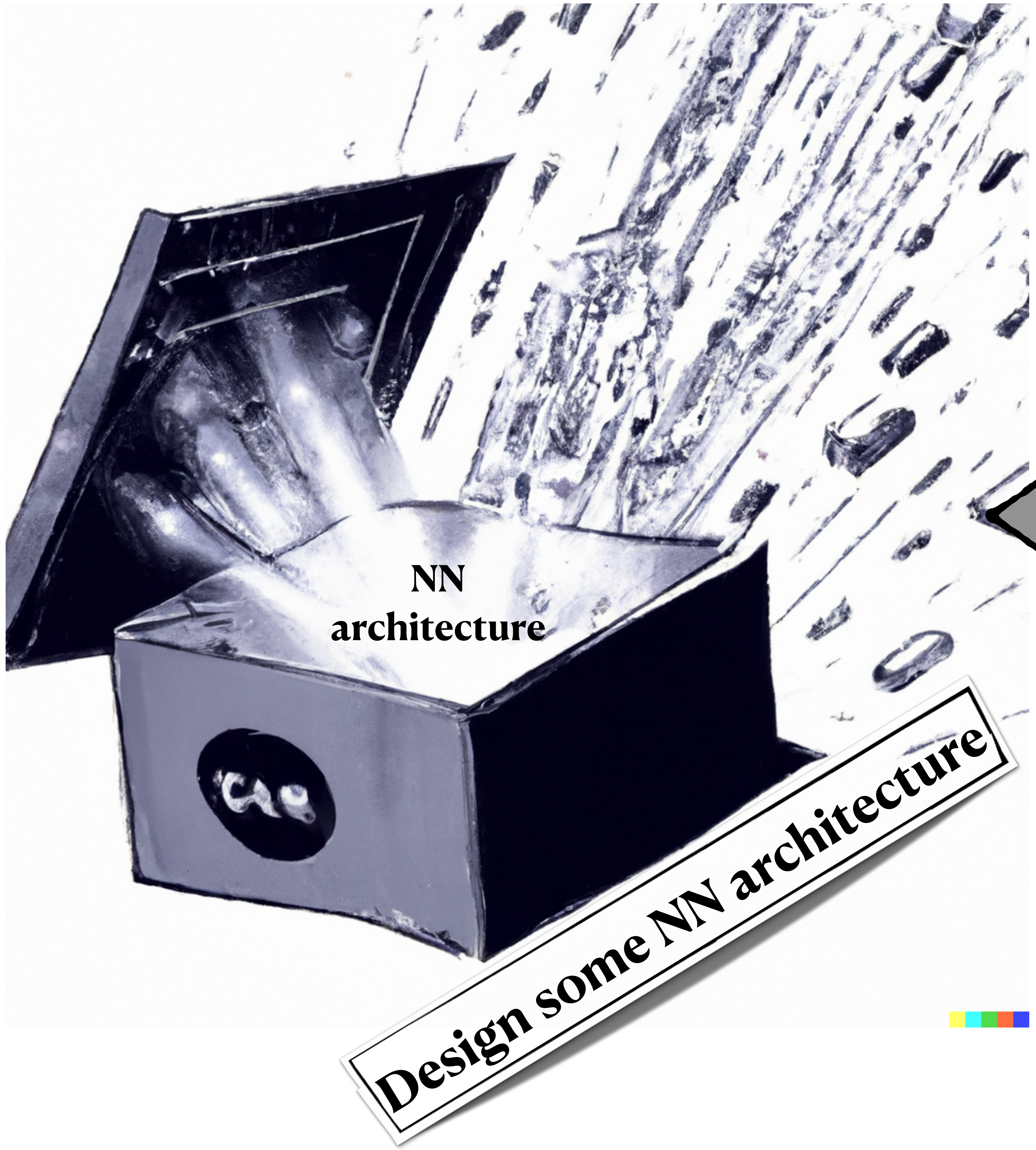
Infinite width N, independently, identically distributed NN parameters

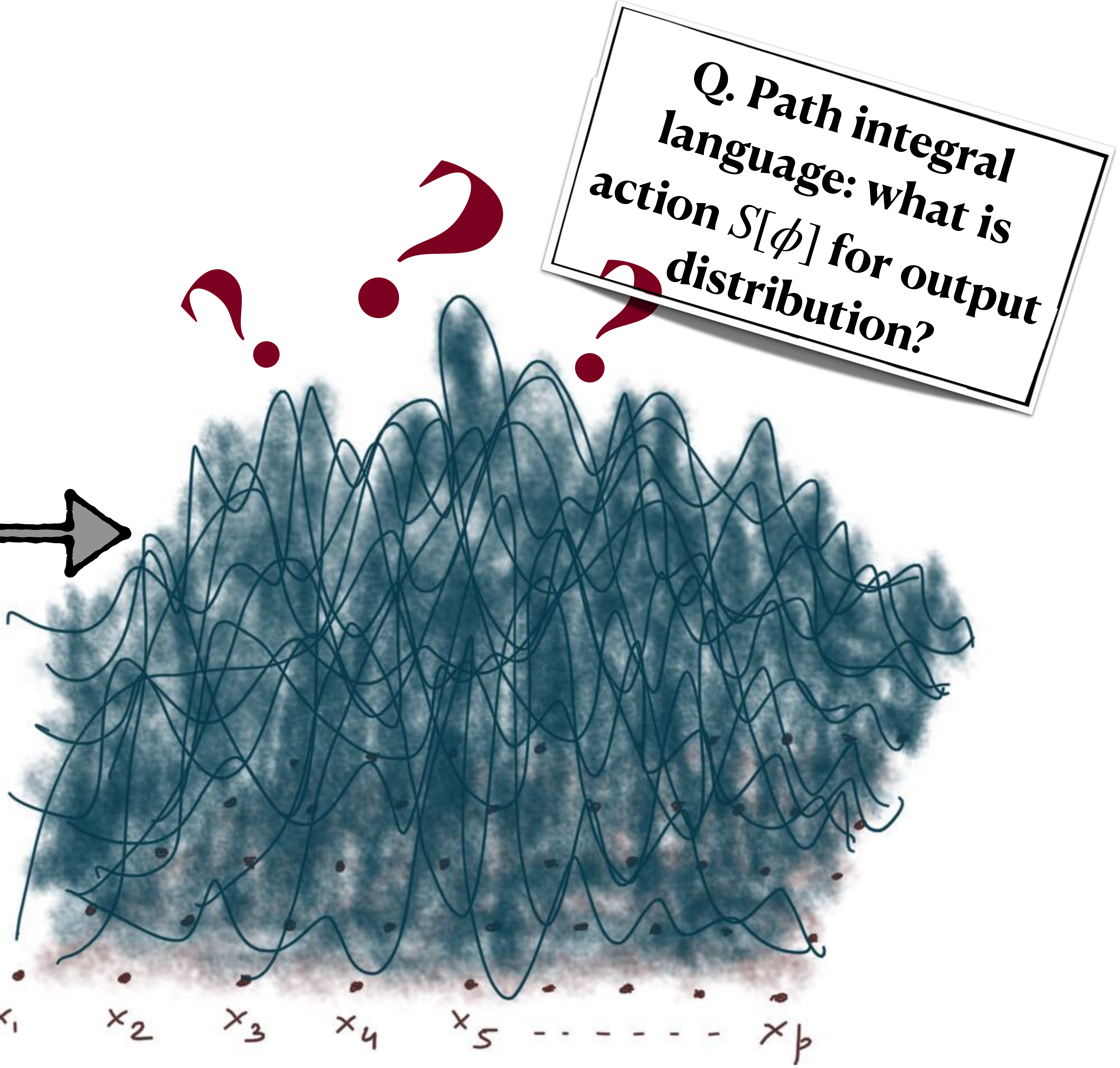
finite width N

Conditions for Central Limit Theorem

- Central Limit Theorem (CLT) constrains NN field distributions as free ones.
- Three axis shows three ways to violate CLT, i.e. turn on field interactions.

NN Field Theory Correspondence





NN Field Action $S[\phi]$

$$Z[J] = \int D\phi \, e^{-S_{\text{free}}[\phi] - S_{\text{int}} + \int d^d x J(x)\phi(x)}$$

Free field action is easy to derive.

- Use NN architecture space to obtain connected parts of n-pt functions.

$$G^{(n)}(x_1, \cdots, x_n) := \mathbb{E}[\phi(x_1)\cdots\phi(x_n)] = \int dh P(h)\phi(x_1)\cdots\phi(x_n)$$

Then,

$$S_{\text{free}}[\phi] = \frac{1}{2}\int d^d x_1 \, d^d x_2 \, \phi(x_1) \, G_c^{(2)}(x_1, x_2)^{-1} \phi(x_2)$$

NN Field Action $S[\phi]$

How to construct couplings $g_r(x_1, \cdots, x_r)$ in $S_{\text{int}} = \sum_{r=3}^{\infty} \int \prod_{i=1}^{r} d^d x_i \, g_r(x_1, \ldots, x_r) \, \phi(x_1) \ldots \phi(x_r)$?

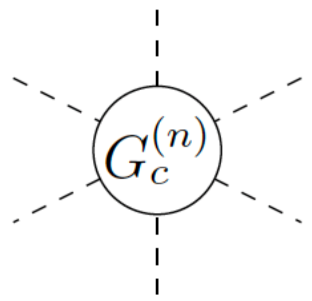
A new set of Feynman rules, to construct $g_r(x_1, \cdots, x_r)$ using connected correlators $G_c^{(r)}$.

Feynman Rules for $g_r(x_1, \ldots, x_r)$.

1. Internal points associated to vertices are unlabeled, for diagrammatic simplicity. Propagators therefore connect to internal points in all possible ways.
2. For each propagator between z_i and z_j,

$$z_i \text{ - - - - } z_j \; = G_c^{(2)}(z_i, z_j)^{-1}. \tag{3.22}$$

3. For each vertex,



$$= (-1)^n \int d^d y_1 \cdots d^d y_n \, G_c^{(n)}(y_1, \cdots, y_n). \tag{3.23}$$

4. Divide by symmetry factor and insert overall $(-)$.

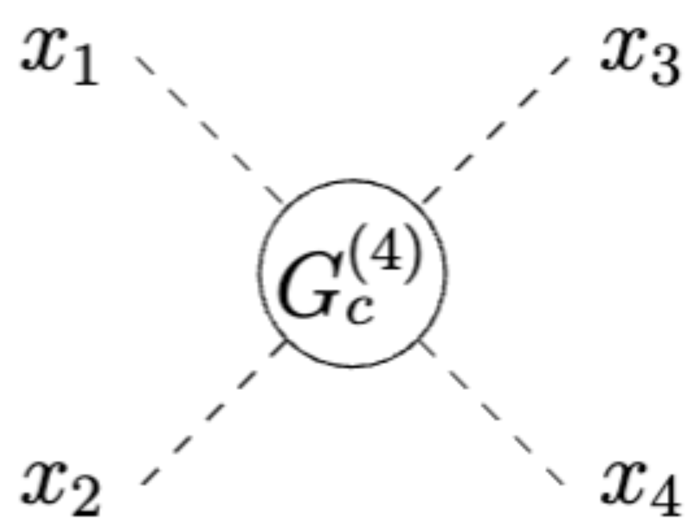
Table 2: Feynman rules for computing g_r from each connected diagram with $G_c^{(n)}$ vertices.

Example: Finite N, i.i.d. parameters

Quartic coupling

$$g_4(x_1,\dots,x_4) = \frac{1}{4!}\Bigg[\int dy_1dy_2dy_3dy_4\, G_c^{(4)}(y_1,y_2,y_3,y_4)\, G_c^{(2)}(y_1,x_1)^{-1}G_c^{(2)}(y_2,x_2)^{-1} G_c^{(2)}(y_3,x_3)^{-1}G_c^{(2)}(y_4,x_4)^{-1} + \text{Comb.}\Bigg] + O\left(\frac{1}{N^2}\right),$$

$$=$$

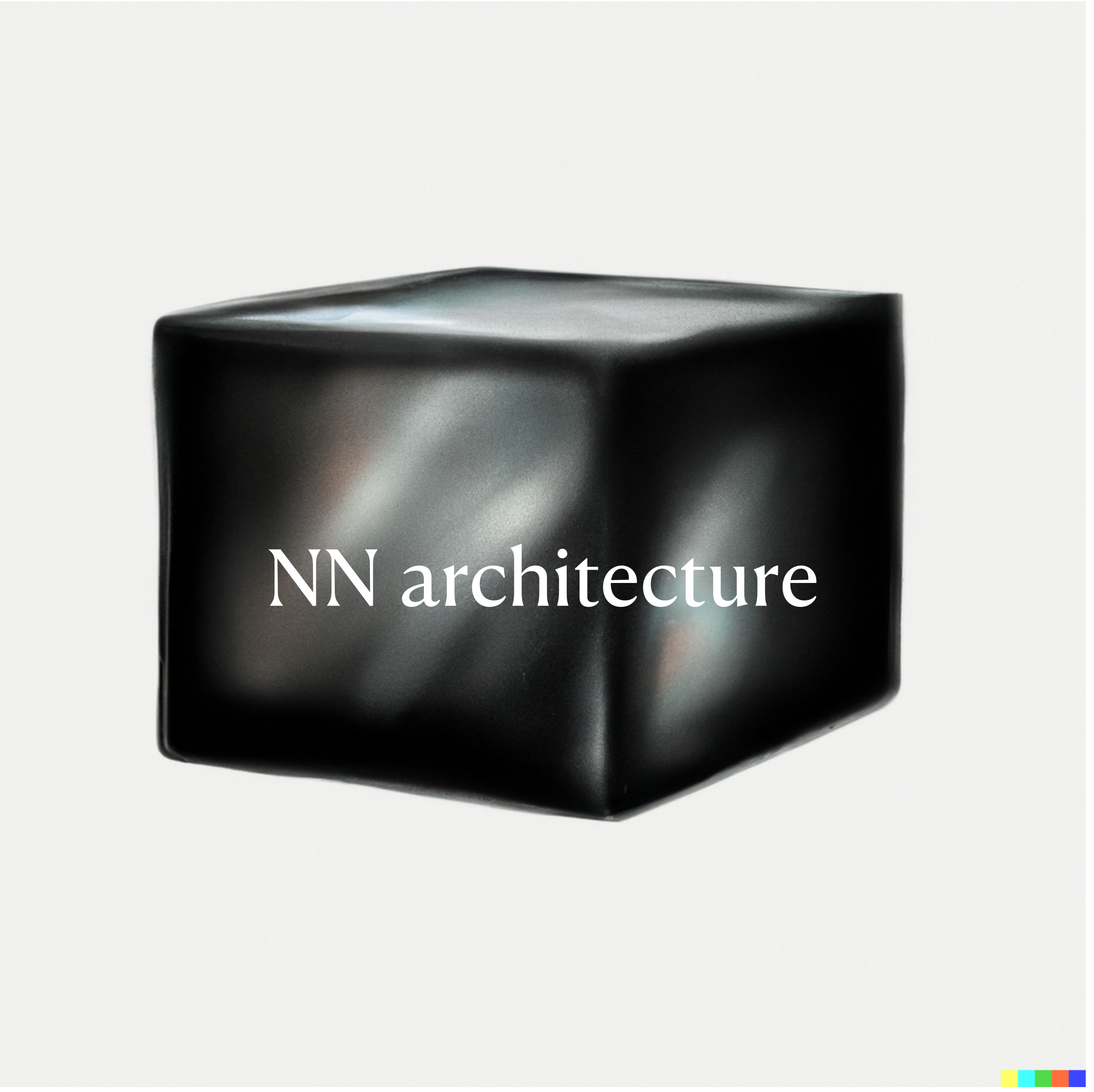


$$+\, O\left(\frac{1}{N^2}\right).$$

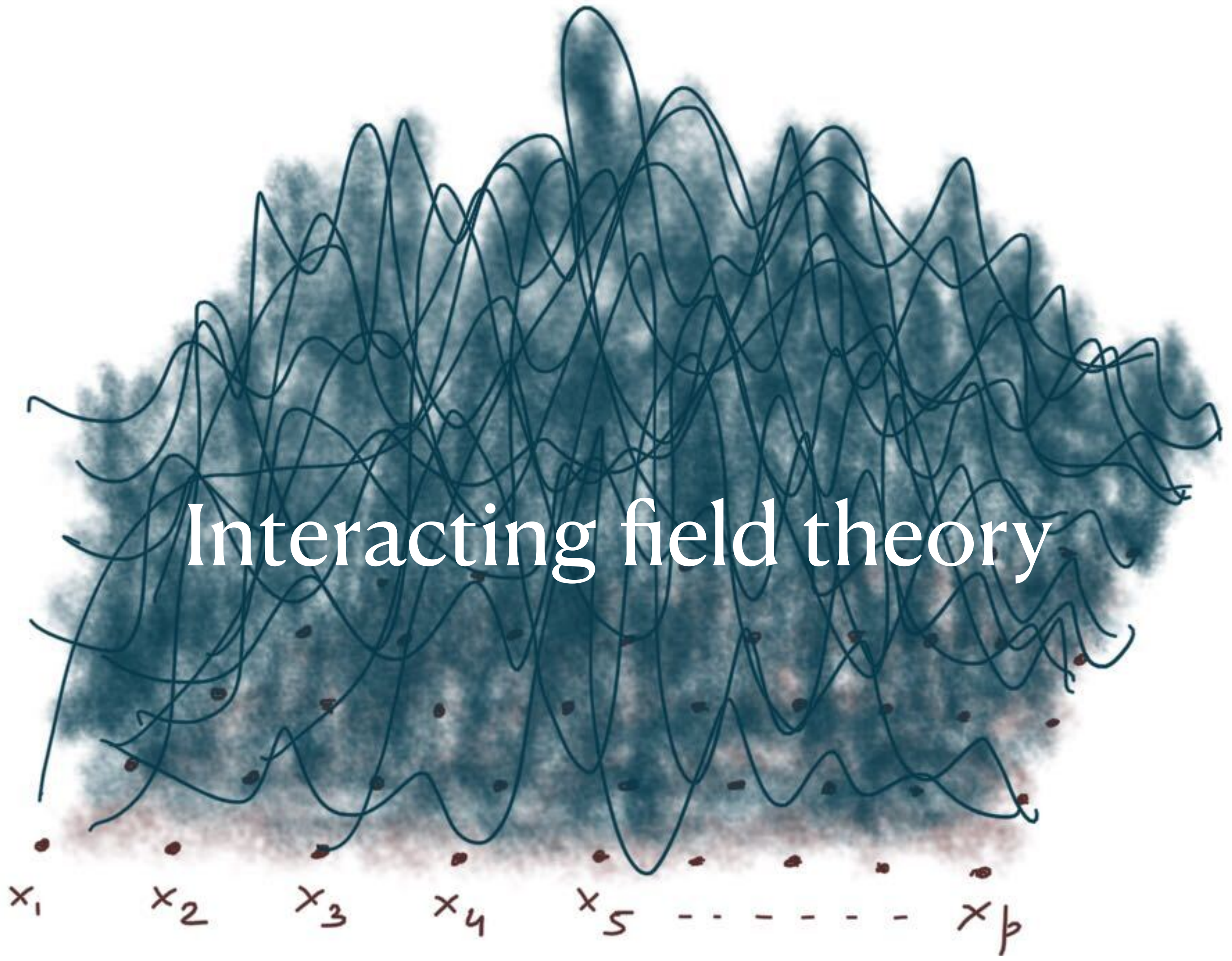
Interacting NN field theory action

$$S[\phi] = S_{\text{free}}[\phi] - \int d^dx_1 \cdots d^dx_4\, g_4(x_1,\cdots,x_4)\,\phi(x_1)\cdots\phi(x_4) + O\left(\frac{1}{N^2}\right)$$

Constrain NN for Given $S[\phi]$



To get field interactions of our choice



$$Z[J] = \int D\phi \, e^{-S[\phi]+\int d^dx J(x)\phi(x)}$$

How to constrain the architecture?

$$S[\phi] = S_{\text{free}}[\phi] + \lambda \int d^dx_1 \cdots d^dx_r \, \mathcal{O}_\phi(x_1, \cdots, x_r)$$

Constrain NN for Given $S[\phi]$

- Getting correct free theory action $S_{\text{free}}[\phi]$ is easy.

◉ $\lim N \to \infty$

◉ i.i.d. parameters $P(h) := P_G(h)$

- Next, deform NN parameters at infinite N,

◉ Redefine $P(h) := P_G(h)\, e^{-\lambda \int d^d x_1 \cdots d^d x_r \, \mathcal{O}_{\phi_h}(x_1, \cdots, x_r)}$

to insert $S_{\text{int}}[\phi] = \lambda \int d^d x_1 \cdots d^d x_r \mathcal{O}_\phi(x_1, \cdots, x_r)$ in NN ensemble action.

Scalar $\lambda\phi^4$ NN field theory

$$S[\phi] = \int d^d x \left[\phi(x)(\nabla^2 + m^2)\phi(x) + \frac{\lambda}{4!}\phi(x)^4 \right]$$

$$\nabla^2 := \frac{\partial^2}{\partial x^2}$$

$$Z[J] = \int da\, db\, dc \;\; P(a,b,c)\; e^{\int d^d x J(x)\, \phi_{a,b,c}(x)}$$

The architecture

$$\phi_{a,b,c}(x) = \sqrt{\frac{2\,\mathrm{vol}(B^d_\Lambda)}{\sigma_a^2 (2\pi)^d}} \sum_{i,j} \frac{a_i \, \cos(b_{ij} x_j + c_i)}{\sqrt{\mathbf{b}_i^2 + m^2}}$$

$\lim N \to \infty$

$$P_G(a) = \prod_i e^{-\frac{N}{2\sigma_a^2} a_i a_i}$$

$$P_G(b) = \prod_i P_G(\mathbf{b}_i) \;\; \text{with} \;\; P_G(\mathbf{b}_i) = \mathrm{Unif}(B^d_\Lambda)$$

$$P_G(c) = \prod_i P_G(c_i) \;\; \text{with} \;\; P_G(c_i) = \mathrm{Unif}([-\pi, \pi])$$

$$P(a,b,c) = P_G(a) P_G(b) P_G(c) \; e^{-\frac{\lambda}{4!}\int d^d x\, \phi_{a,b,c}(x)^4}$$

Conclusion & Outlook

- NN output ensembles are field distributions.
- CLT constrains NN field theory as a free one. Violation of CLT turns on interaction terms.
- NNFT action $S[\phi]$ can be obtained using new Feynman rules.
- Some interacting NNFT are obtained by appropriately distorting NN parameter distributions.
- E.g. an architecture for $\lambda\phi^4$ scalar field theory is obtained at inf N, via parameter deformations from i.i.d.
- Currently, we are working on NN architectures at initialization for fermionic field theories.

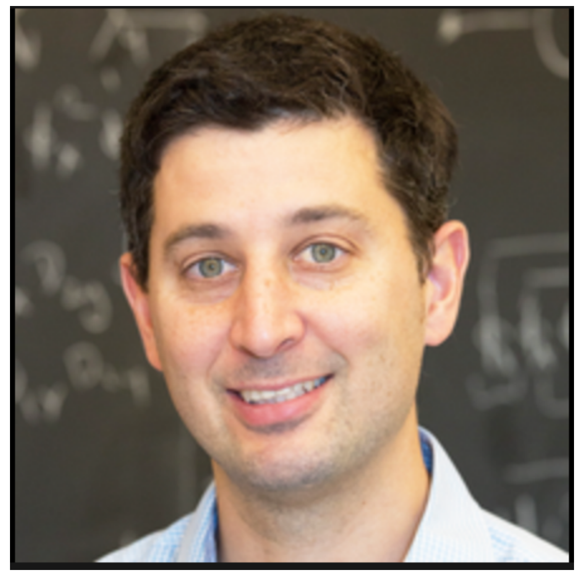
My amazing collaborators!

Thank you!

Questions?

Feel free to get in touch:
email: amaiti@perimeterinstitute.ca
Twitter: @AninditaMaiti7
web: https://aninditamaiti.github.io/

Jim Halverson



Matt Schwartz

Mehmet Demirtas

Keegan Stoner