A Neural Network Approach to Lattice Field Theory

Andy Sheng[1]

Lattice Conference 2023
Fermilab, Batavia, IL
7/31/2023

Paulo Bedaque[1]
Andrea Carosso[2]

[1]University of Maryland College Park

[2]George Washington University

Motivation

- Quantum computation → Hamiltonian formalism of QFTs (Schrödinger picture)
- Hilbert space is exponentially large
- Spin system Hamiltonians: tensor networks (MPS) describe ground state/excited state wavefunctions using polynomially many numbers
- Bosonic field theories: truncate to “spin system” to use MPS
- Use neural networks to approximate the un-truncated theory?
- Encode ground state/excited wave-functional properties of QFTs using a (polynomial?) number of parameters
- More efficient way of simulating imaginary-time bosonic QFTs?

Schrödinger Picture QFT

- Hamiltonian operator:

$$H = a^d \sum_x \left[\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}\sum_{k=1}^{d}\left(\frac{\hat{\phi}_{x+k} - \hat{\phi}_x}{a}\right)^2 + \frac{1}{2}m^2\hat{\phi}_x^2 + V(\hat{\phi}_x) \right]$$

- Satisfy canonical commutation relations $[\hat{\phi}_x, \hat{\pi}_y] = ia^{-d}\delta_{x,y}$
  - Choose $\hat{\phi}_x = \phi_x, \ \hat{\pi}_x = -ia^{-d}\frac{\partial}{\partial \phi_x}$

$$H = a^d \sum_x \left[-\frac{1}{2a^{2d}}\frac{\partial^2}{\partial \phi_x^2} + \frac{1}{2}\sum_{k=1}^{d}\left(\frac{\phi_{x+k} - \phi_x}{a}\right)^2 + \frac{1}{2}m^2\phi_x^2 + V(\phi_x) \right]$$

- 1D non-relativistic many-body Hamiltonian
- Ground state wave-functional $\psi(\Phi) = \psi(\phi_1, \ldots, \phi_N)$

Variational Monte Carlo

Variational Ansatz: $\psi = \psi(\Phi, \theta)$

i.e. $\psi(\Phi, \theta) \sim e^{\theta_1 \phi^2 + \theta_2 \phi^4 + \theta_3 \phi^6 + \dots}, \quad \theta = \{\theta_1, \theta_2, \dots\}$

$$\langle E \rangle = \frac{\int d\Phi \, \psi^* \psi \, \psi^{-1} H \psi}{\int d\Phi \, \psi^* \psi}$$

Estimate using Monte Carlo (Metropolis, etc.)

$$\frac{\partial \langle E \rangle}{\partial \theta} = 2 \frac{\int d\Phi \, \psi^* \psi \, \psi^{-1} (\partial \log \psi / \partial \theta) H \psi}{\int d\Phi \, \psi^* \psi} - 2 \langle E \rangle \frac{\int dX\Phi \, \psi^* \psi (\partial \log \psi / \partial \theta)}{\int d\Phi \, \psi^* \psi}$$

$$\langle E \rangle = \frac{\langle \psi(\Phi, \theta) \,|\, H \,|\, \psi(\Phi, \theta) \rangle}{\langle \psi(\Phi, \theta) \,|\, \psi(\Phi, \theta) \rangle} \geq \langle E \rangle_0$$

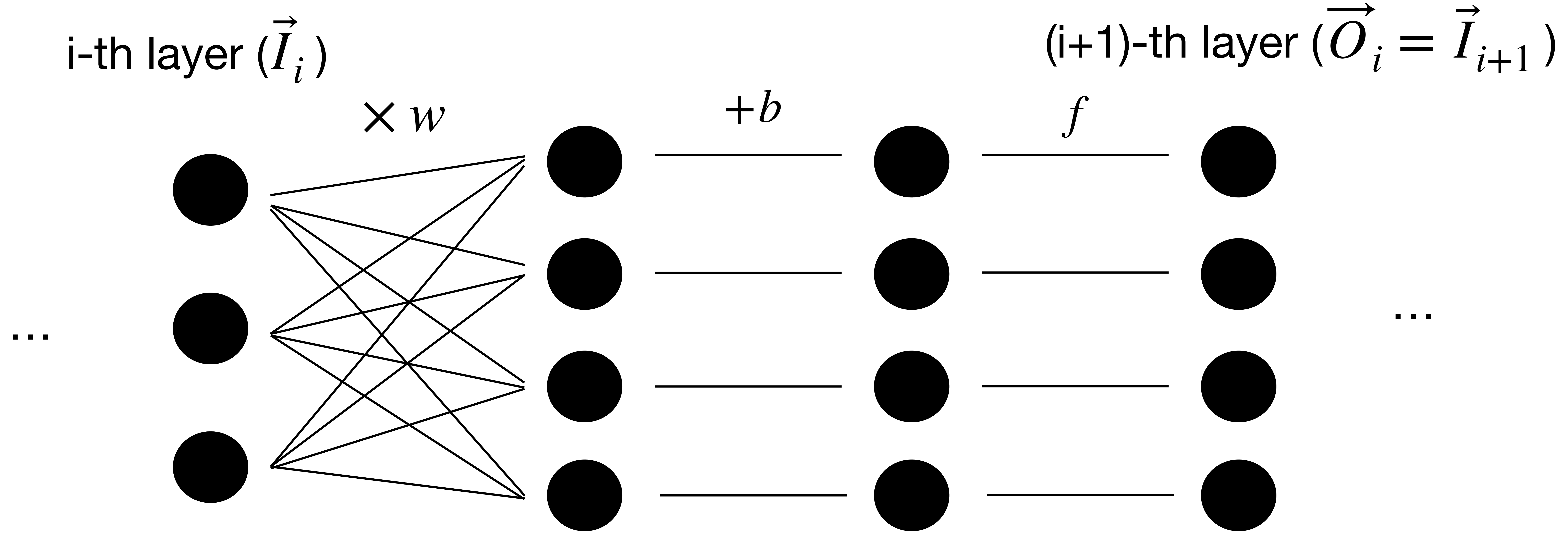
- Gradient descent to minimize $\langle E \rangle$ with respect to θ
- $\min_\theta \left[\langle E(\theta) \rangle \right]$ provides upper bound for $\langle E \rangle_0$
- Accuracy of estimate:
  - Number of parameters in the ansatz → systematically increase?
  - Functional form of $\psi(\Phi, \theta)$ → generality?
- Use neural network functions…

Neural Networks

Function mapping N inputs to M outputs (M = 1)

$$\vec{I}_{i+1} = \vec{O}_i = f(\mathbf{W}_i \vec{I}_i + \vec{b}_i)$$

Non-linear activation function — Weights — Biases



- Universal function approximator (Hornik et. al. 1989, Hornik 1991)

Neural Network Ansatze

- $\psi \sim \mathcal{N}(\Phi, \theta), \quad \theta = \{\mathbf{W}_i, \overrightarrow{b}_i\},$ Easily increase number of parameters
- Ground states of bosonic theories are real + positive
  - $\psi \sim \psi_{\text{free}} \cdot e^{-\mathcal{N}(\Phi,\theta)}$
  - Neural network encodes information about interactions
- Enforce symmetries of the theory (i.e Bose exchange, translation invariance, …)
- $\langle E \rangle \to$ cost function, gradient step $\to$ training step
- Can compute $\dfrac{\partial \mathcal{N}}{\partial \theta}$ easily using backpropagation
- Standard gradient descent algorithms (i.e Adam, …)

Proof of Principle

Interacting bosons via short and long-range potentials (Beau et. al. 2020)

$$H = \sum_{i=1}^{N} -\frac{1}{2}\frac{\partial}{\partial x_i^2} + \frac{1}{2}x_i^2 + \sum_{i<j} g\delta(x_i - x_j) - \frac{g}{2}|x_i - x_j|$$

“Free theory”: Kinetic + Harmonic Trap

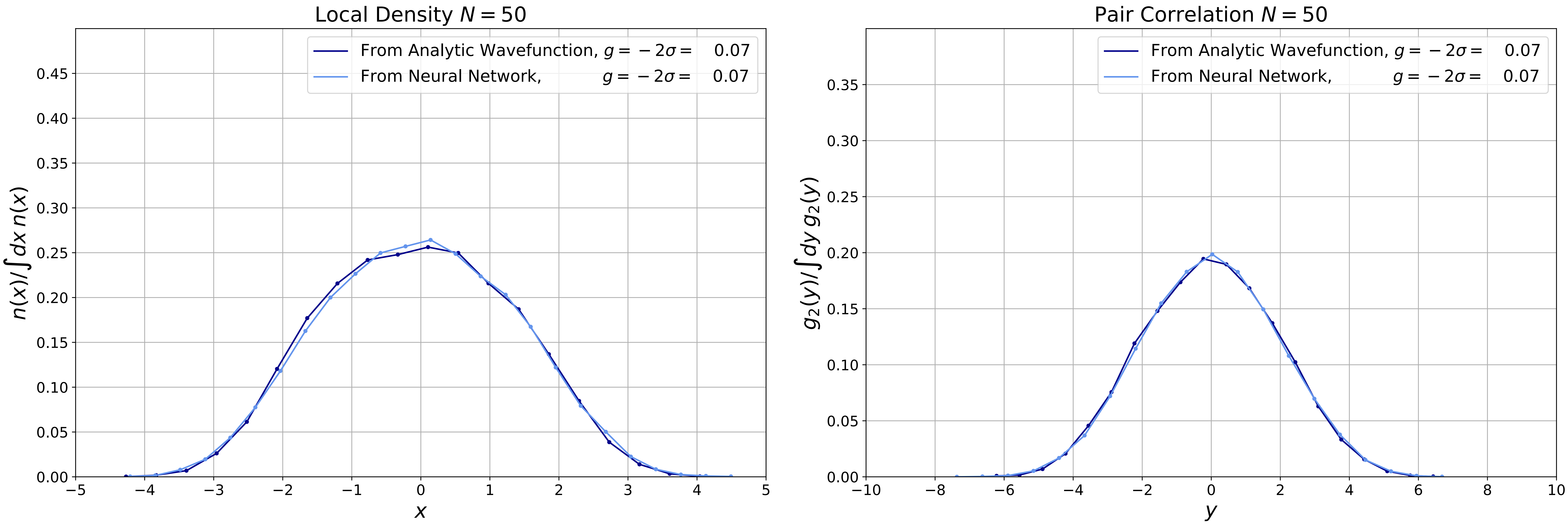
Potential: Contact
+ Long-range linear

- Exactly solvable in one dimension
- Rich variety of “phases” depending on sign of g

Ground state behavior $g > 0$

Local density: $n(x) = \int dx_2 \ldots dx_N \, \psi_0(x = x_1, x_2, \ldots, x_N)^2$

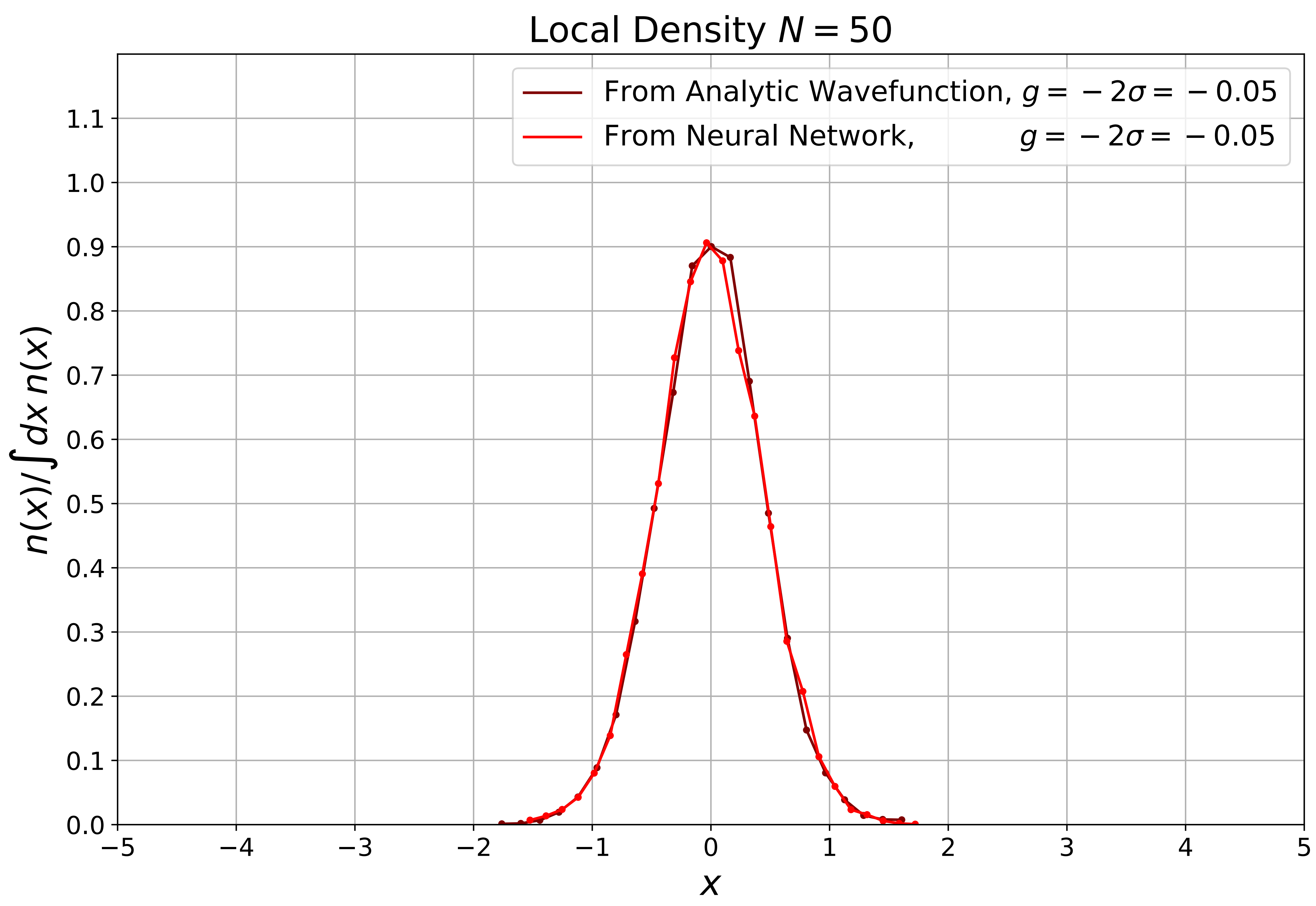
Pair density: $n(y) = \int dX \, \delta(y - |x_1 - x_2|) \psi_0(X)^2$

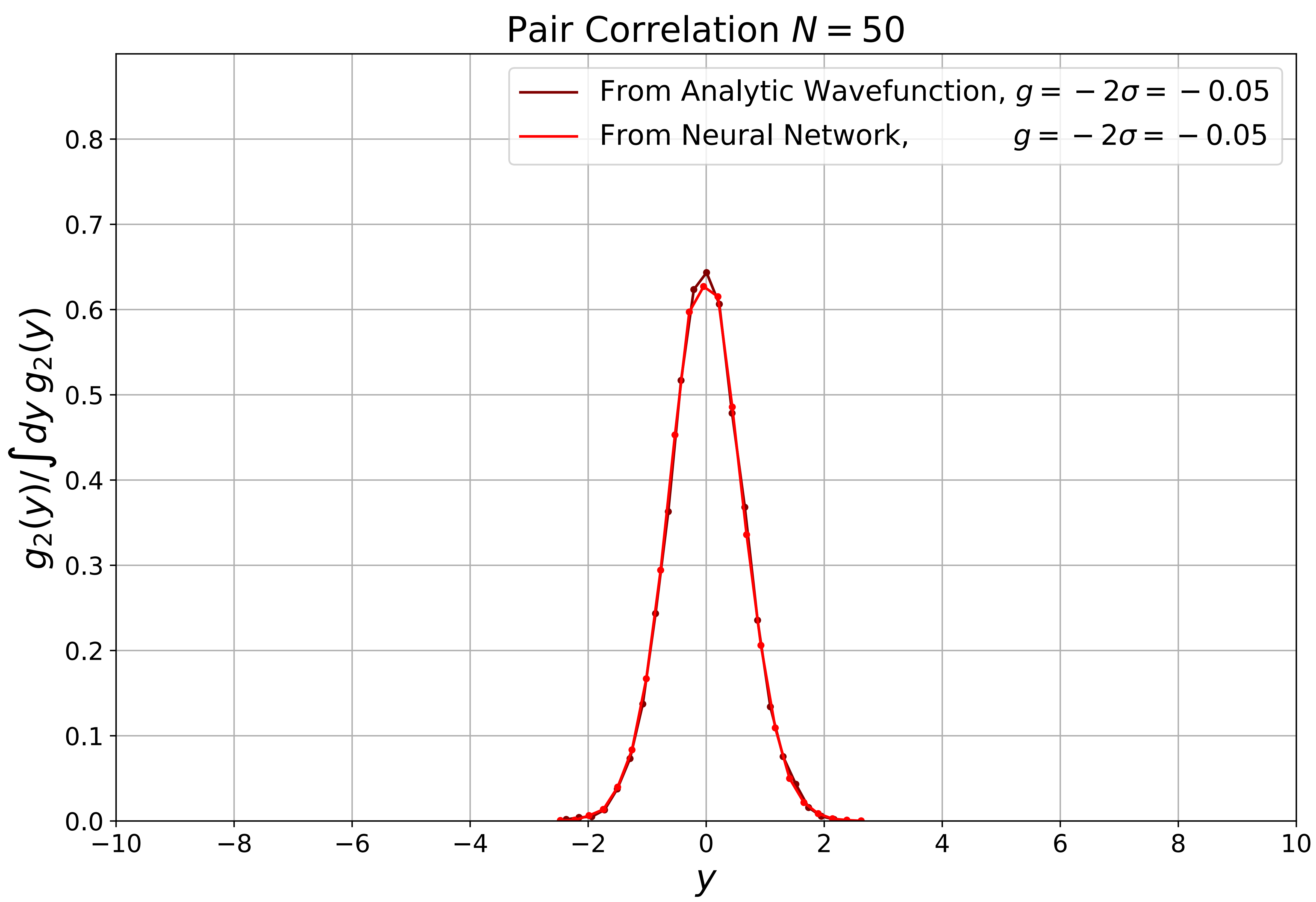


Ground state behavior $g < 0$

Local density: $n(x) = \int dx_2 \ldots dx_N \, \psi_0(x = x_1, x_2, \ldots, x_N)^2$

Pair density: $n(y) = \int dX \, \delta(y - |x_1 - x_2|)\psi_0(X)^2$





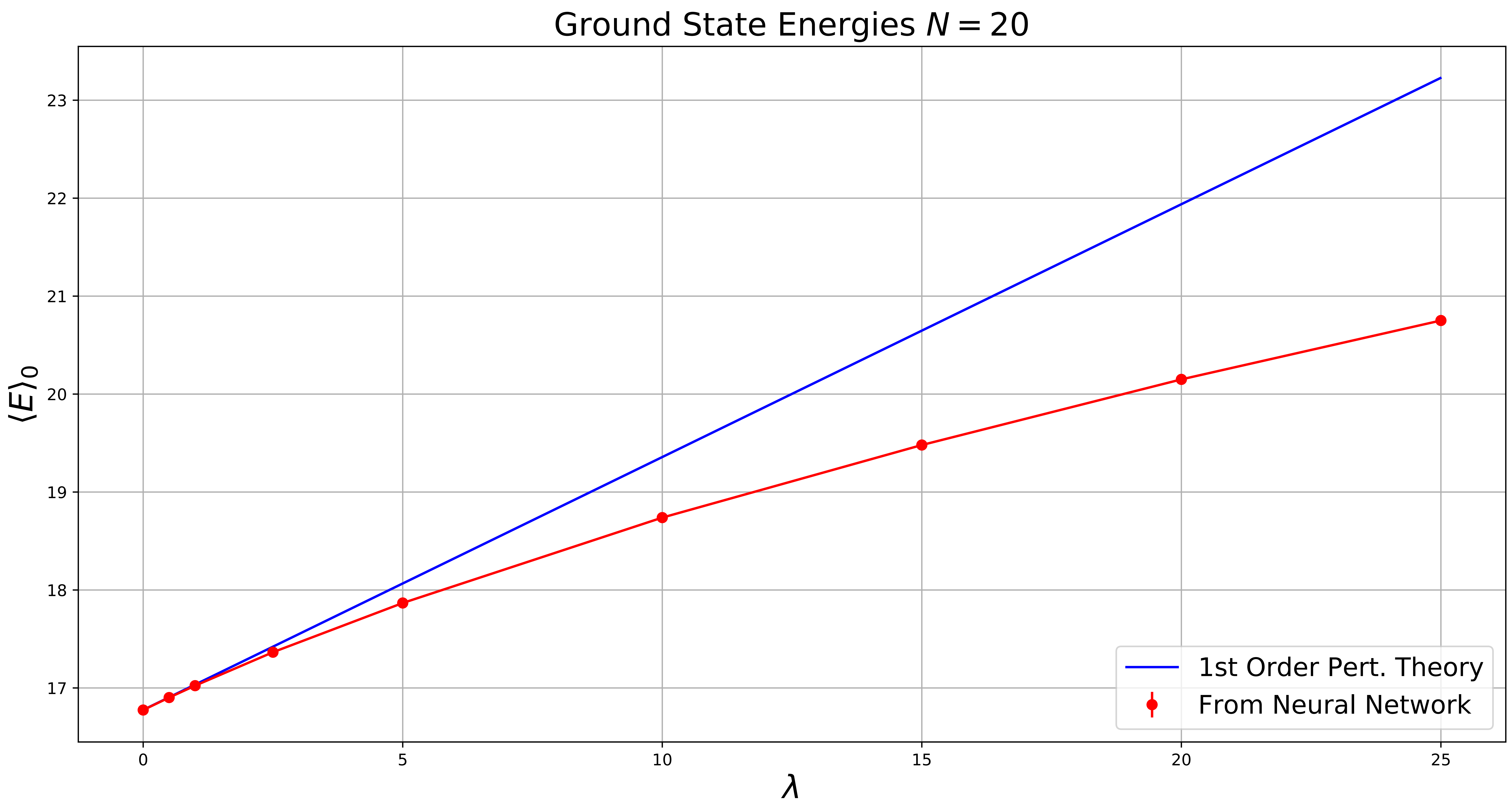
ϕ^4-Theory: Preliminary Results

$$H = a_x \sum_{x=1}^{N_x} \left[\underbrace{-\frac{1}{2a_x^2}\frac{\partial^2}{\partial \phi_x^2} + \frac{1}{2}\left(\frac{\phi_{x+1} - \phi_x}{a_x}\right)^2}_{H_0 \text{ (free)}} + \underbrace{\frac{1}{2}m^2\phi_x^2 + \frac{\lambda}{4!}\phi_x^4}_{V} \right]$$

- Preliminarily study in 1+1 dimensions
- Periodic boundary conditions $\phi_{x+N} = \phi_x$,
- Fix $a_x = 1,\ m = 1$
- $\psi(\phi_1, .., \phi_{N_x}, \theta) \propto \psi_{\text{free}} \cdot e^{-\underbrace{\left(\mathcal{N}(\phi_1, ..., \phi_N) + \mathcal{N}(\phi_2, ..., \phi_N, \phi_1) + ...\right)}_{\text{Enforces translation invariance}}}$
- $\vec{\phi} \to -\vec{\phi}$ symmetry not enforced, learned?

Ground State Energies

$$E_0^1 = \langle \psi_{\text{free}} | H_0 + V | \psi_{\text{free}} \rangle > \langle \psi_0 | H_0 + V | \psi_0 \rangle$$



Point-point Spatial Correlators

From the Hamiltonian

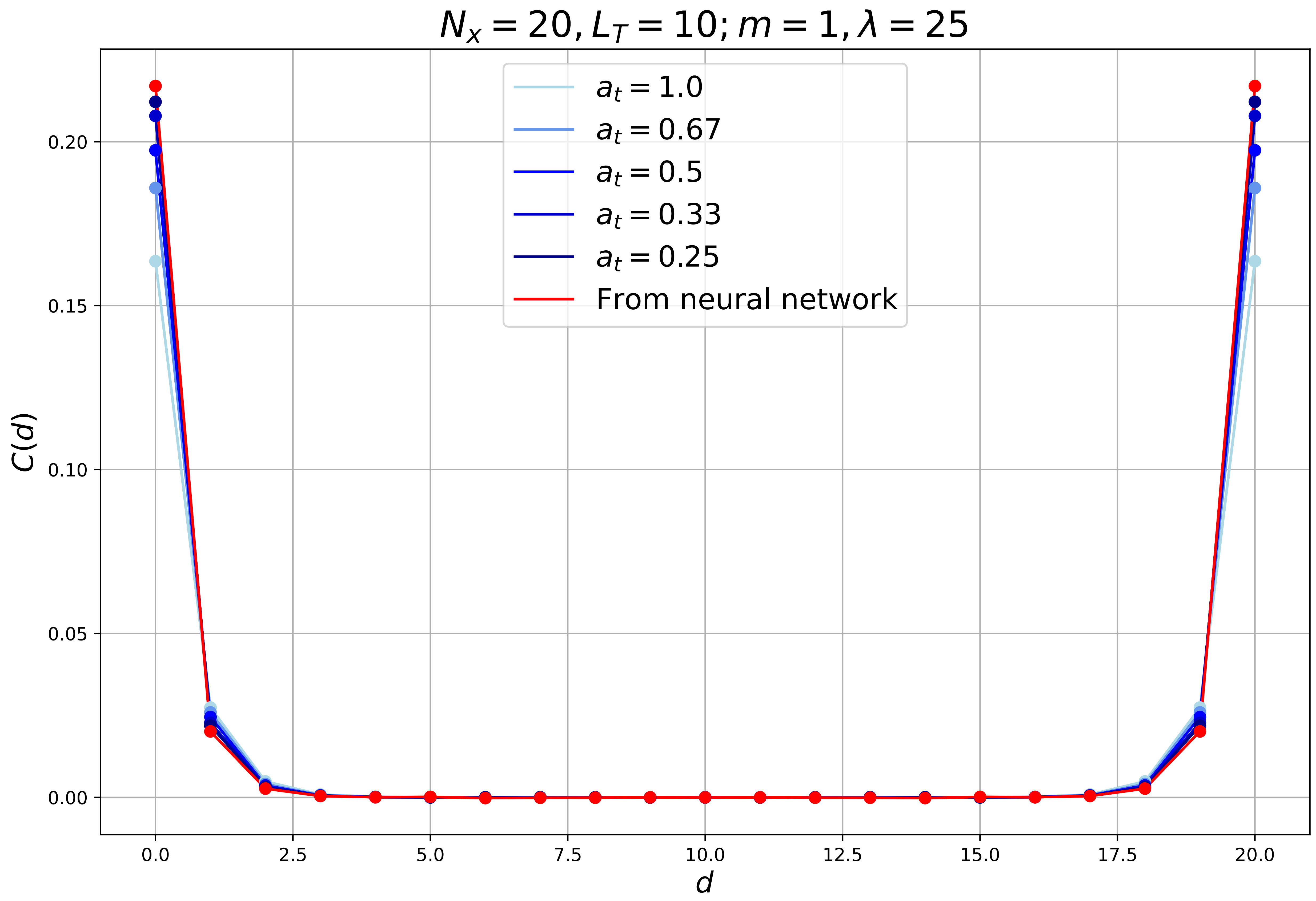
$$C_{NN}(d) = \frac{1}{N_x} \sum_{x=1}^{N_x} \langle \phi_x \phi_{x+d} \rangle$$

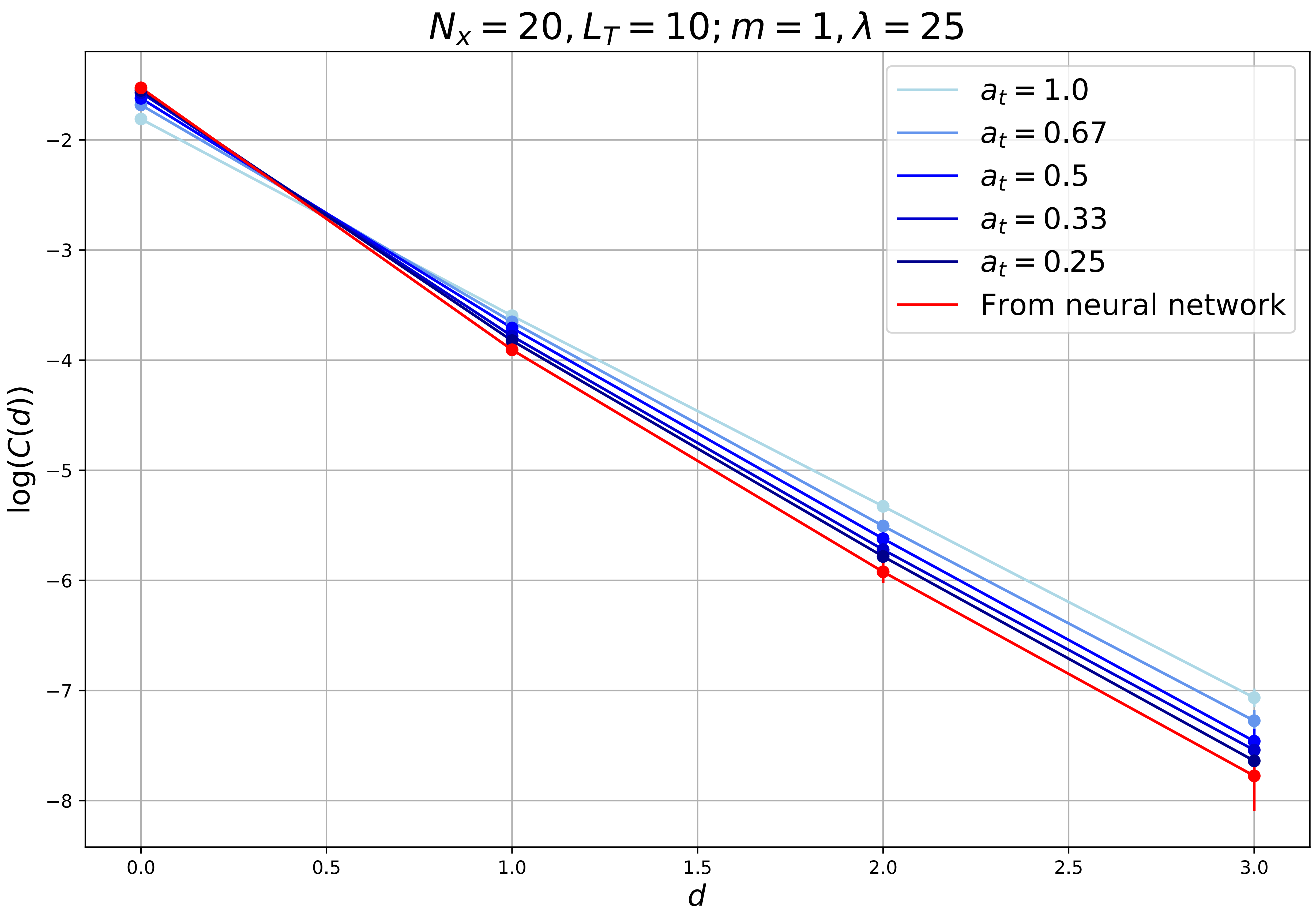
From the Path Integral:

$$C_{PI}(d) = \frac{1}{N_t N_x} \sum_{t,x}^{N_t,N_x} \langle \phi_{t,x} \phi_{t,x+d} \rangle$$

- $C_{PI}(d)$ should approach $C_{NN}(d)$ in time-continuous limit: $a_t \to 0$
- Hamiltonian: in the zero-temperature limit $- \, L_T = N_t a_t \to \infty$
- Finite temporal extent effects $\sim e^{-amN_t}$ where am is the inverse spatial correlation length
    - Effects negligible if we fix $amN_t \gtrsim 8 \quad (e^{-8} \approx 10^{-4})$

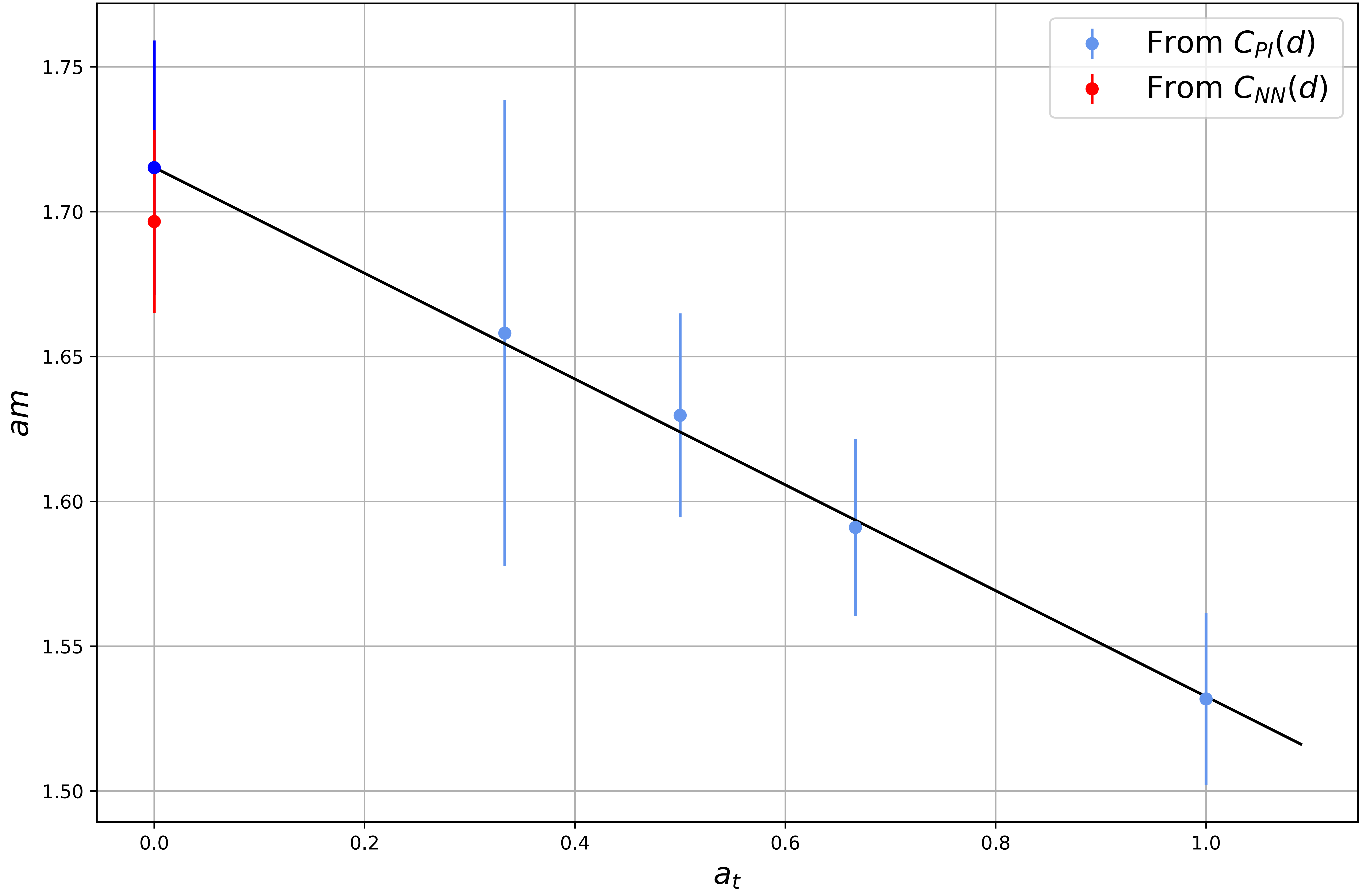
Point-point Spatial Correlators





Spatial Correlation Lengths

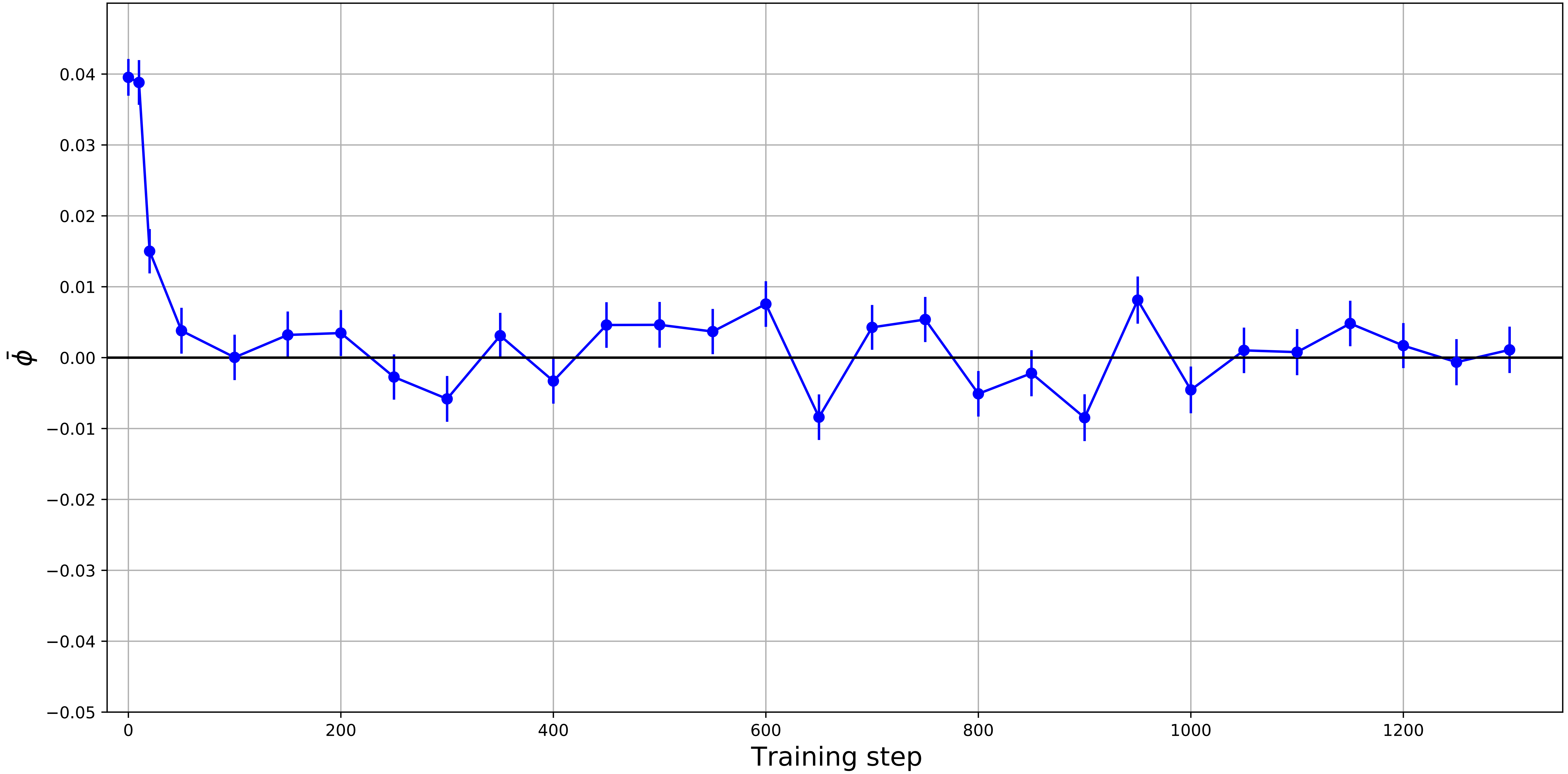
- Extract am by fitting to $C(d) = A\left[K_0(am \cdot d) + K_0(am \cdot (N_x - d))\right]$



"Learning" Z_2 Symmetry

- ϕ^4-theory is symmetric under $\phi \to -\phi$
- Ground state must respect the symmetry

$$\bar{\phi} = \frac{1}{N_x} \sum_{i=1}^{N_x} \langle \phi_i \rangle$$



Conclusions

- Schrödinger picture scalar QFT on a lattice → 1D non-relativistic many-body quantum mechanics
- Solve using Variational Monte Carlo with neural network wavefunction ansatze

Future Studies

- Scaling: number of NN parameters vs. number of lattice sites
- Quantify the cost of the neural network vs. Monte Carlo
- Efficiently and effectively take the continuum limit $a_x \to 0$
  - Avoid critical slowing down?
- Increase the number of dimensions to 2+1, 3+1

Realness and Positivity of Ground States

Realness

$|\psi\rangle = |\psi_R\rangle + i\,|\psi_R\rangle$

$H|\psi_R\rangle = E_0|\psi_R\rangle,\; H|\psi_I\rangle = E_0|\psi_I\rangle$

For a non-degenerate ground state, can choose w.l.o.g $|\psi_I\rangle = 0$

Positivity

$$\langle E\rangle_0 = -\frac{1}{2m}\int dX\,\sum_{i=1}^{N}\psi\frac{\partial^2\psi}{\partial x_i^2} + \int dX\,V(X)\psi^2$$

$$= \frac{1}{2m}\int dX\,\sum_{i=1}^{N}\left(\frac{\partial\psi}{\partial x_i}\right)^2 + \int dX\,V(X)\psi^2$$

If allowed, $|\psi(X)|$ also a ground state; choose w.l.o.g $\psi(X)$ to be positive

Enforcing Bose Symmetry in One Dimension

- Wavefunction of a bosonic system must satisfy $\psi(\ldots, x_i, \ldots, x_j, \ldots) = \psi(\ldots, x_j, \ldots, x_i, \ldots), \forall\, 1 \leq i \neq j \leq N$
- $\psi(X, \theta) \propto$ (free) $\cdot\, e^{-\mathcal{N}(x_1, \ldots, x_N)}$, need to make $\mathcal{N}(x_1, \ldots, x_N)$ satisfy Bose symmetry
- $\{x_1, \ldots, x_N\} \rightarrow \{\xi_1, \ldots, \xi_N\}$ is a bijection for

$$\xi_n = \sum_{i=1}^{N} x_i^n$$

- Use $\{\xi_1, \ldots, \xi_N\}$ as Bose symmetric inputs to the neural network

$$\psi(X, \theta) \propto \text{(free)} \cdot e^{-\mathcal{N}(\xi_1, \ldots, \xi_N)}$$

Neural Network Convergence

- Keep the same number of layers, increase number of nodes/layer
- β: total number of parameters (weights and biases) in the neural network