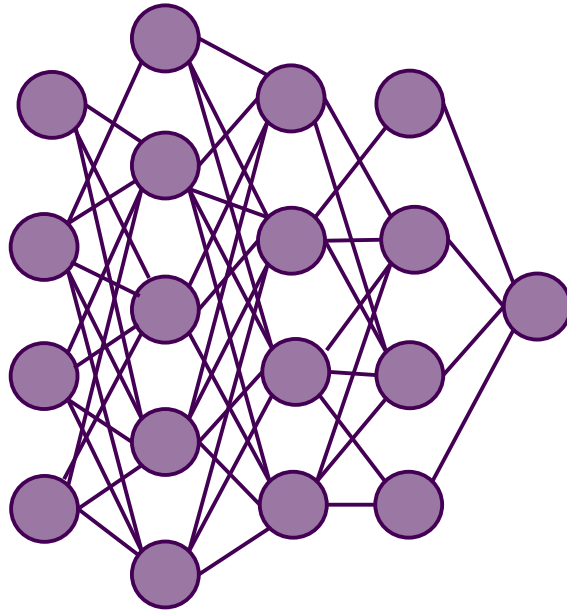


Variational wave-functions and neural networks

Agnes Valenti



Based on:

*Lecture on NQS at ICTP 2024 (smr 3928)
by Filippo Vicentini*

*Review article: M Medvidovic, J Robledo
Moreno, arXiv:2402.11014 (2024)*

*Book: F Becca, "Quantum Monte Carlo
approaches for correlated systems"*

Part I

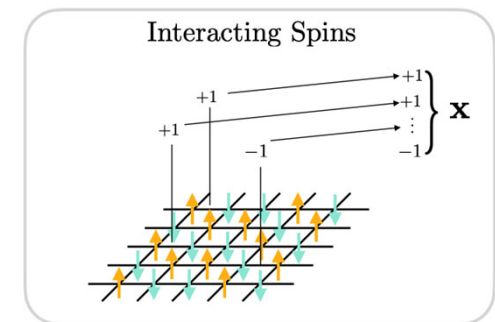
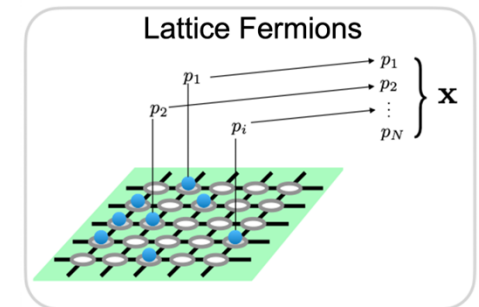
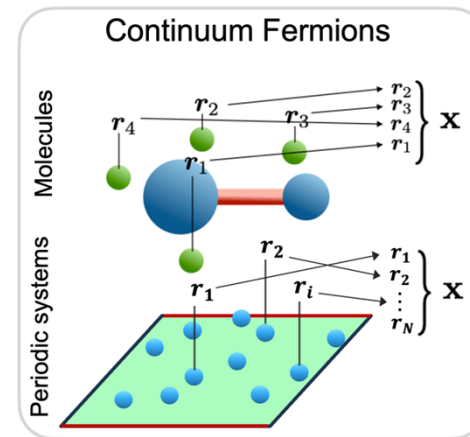
Machine learning for quantum many-body simulations

Quantum many-body problem(s)

Schroedinger equation $\mathcal{H}|\Psi\rangle = i\hbar\partial_t|\Psi\rangle$

Quantum many-body problem(s)

Schroedinger equation $\mathcal{H}|\Psi\rangle = i\hbar\partial_t|\Psi\rangle$



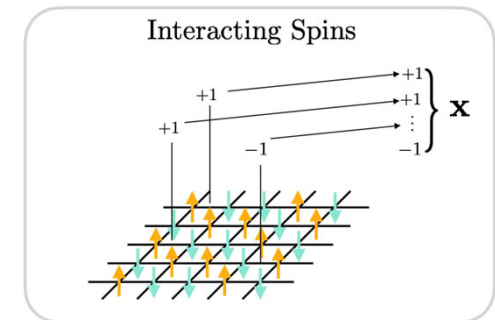
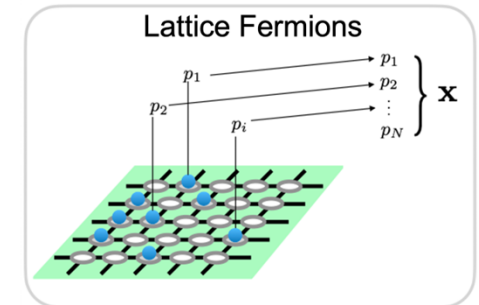
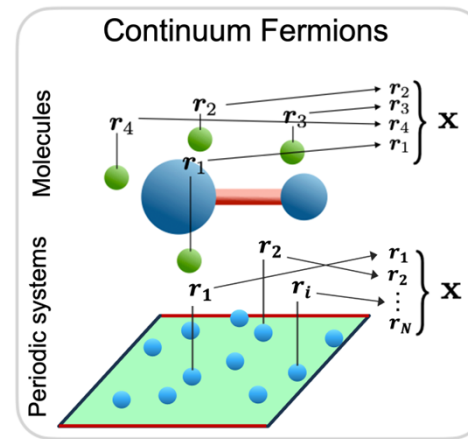
[M Medvidovic, J Robledo Moreno, arXiv:2402.11014 (2024)]

Quantum many-body problem(s)

Schroedinger equation $\mathcal{H}|\Psi\rangle = i\hbar\partial_t|\Psi\rangle$

- Ground states $\mathcal{H}|\Psi\rangle = E|\Psi\rangle$
- Time evolution $|\Psi_0\rangle \rightarrow |\Psi(t)\rangle$
- Finite temperature $[\mathcal{H}, \rho] = i\hbar\partial_t\rho$
- Open systems $\mathcal{L}(\rho) = \partial_t\rho$

⋮



[M Medvidovic, J Robledo Moreno, arXiv:2402.11014 (2024)]

Quantum many-body problem

Find the ground state $\mathcal{H}|\Psi\rangle = E|\Psi\rangle$

Quantum many-body problem

Find the ground state $\mathcal{H}|\Psi\rangle = E|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S)|S\rangle$$

Quantum many-body problem

Find the ground state $\mathcal{H}|\Psi\rangle = E|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S) |S\rangle$$

$S = s_1, \dots, s_N$



Basis state: Spin-configuration

Quantum many-body problem

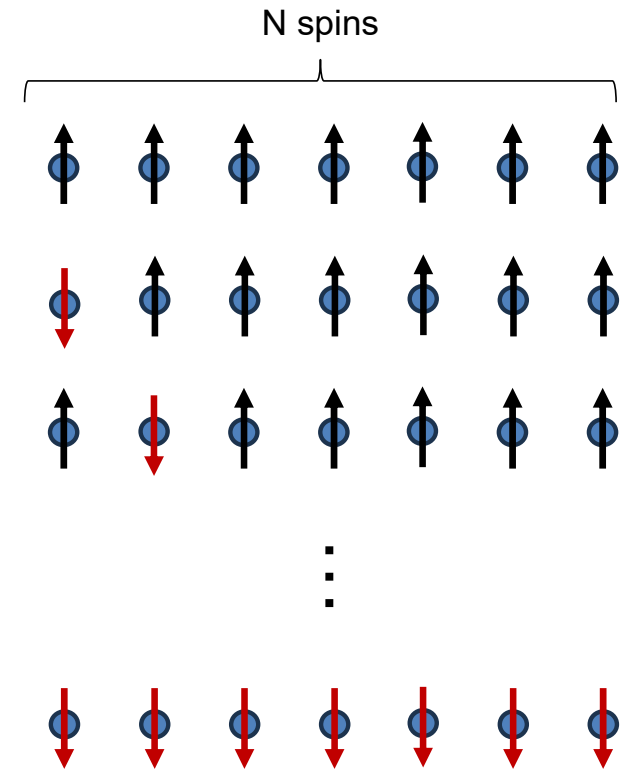
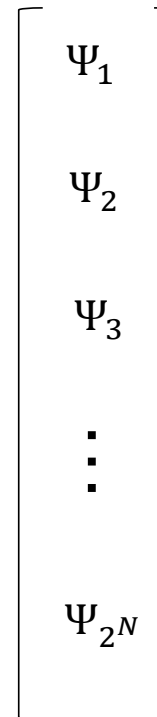
Find the ground state $\mathcal{H}|\Psi\rangle = E|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S) |S\rangle =$$

$S = s_1, \dots, s_N$



Basis state: Spin-configuration

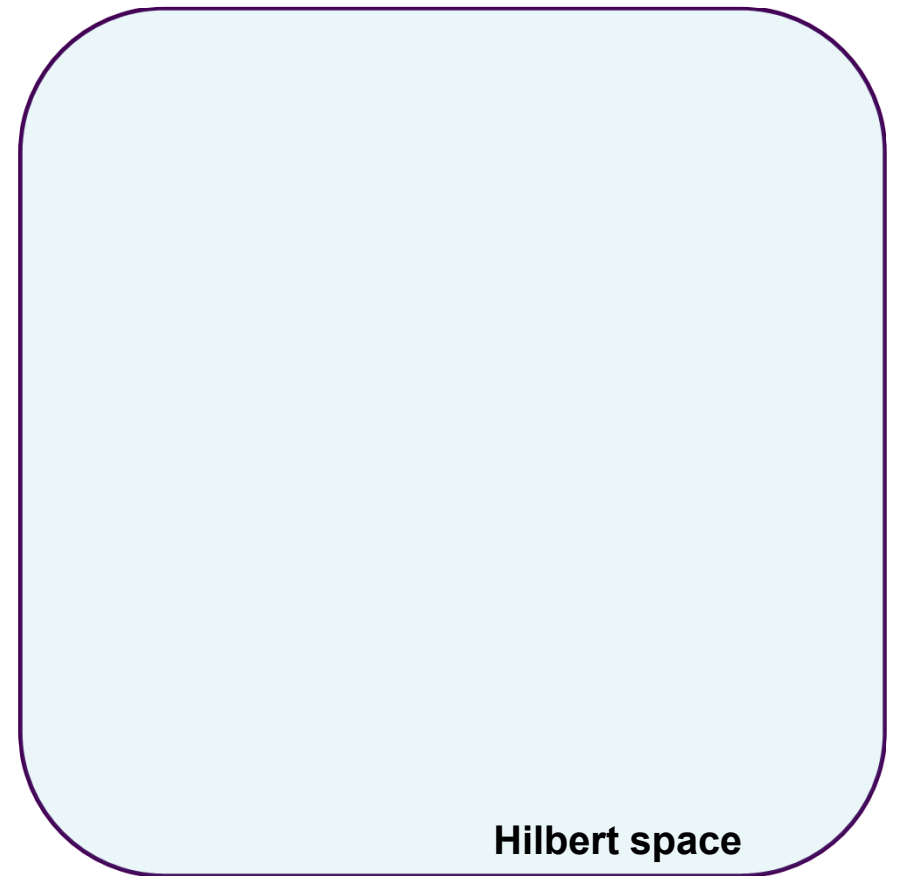


⇒ Hilbert space dimension 2^N

Variational wave-functions

Find the ground state

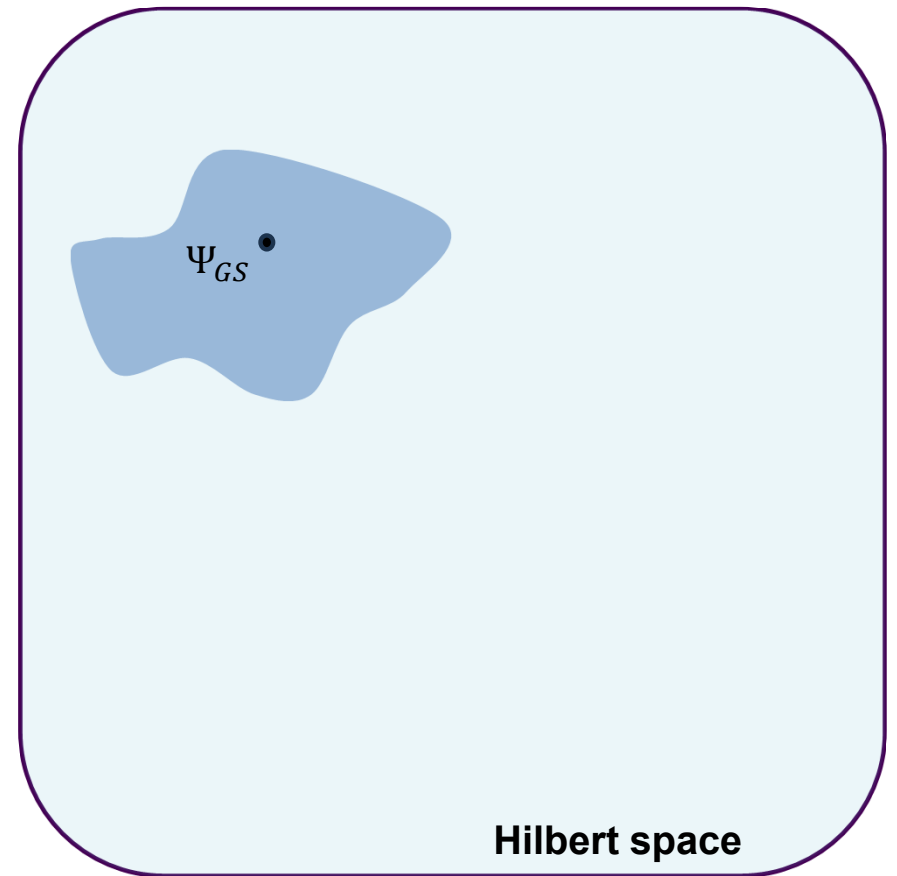
$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$



Variational wave-functions

Find the ground state

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$



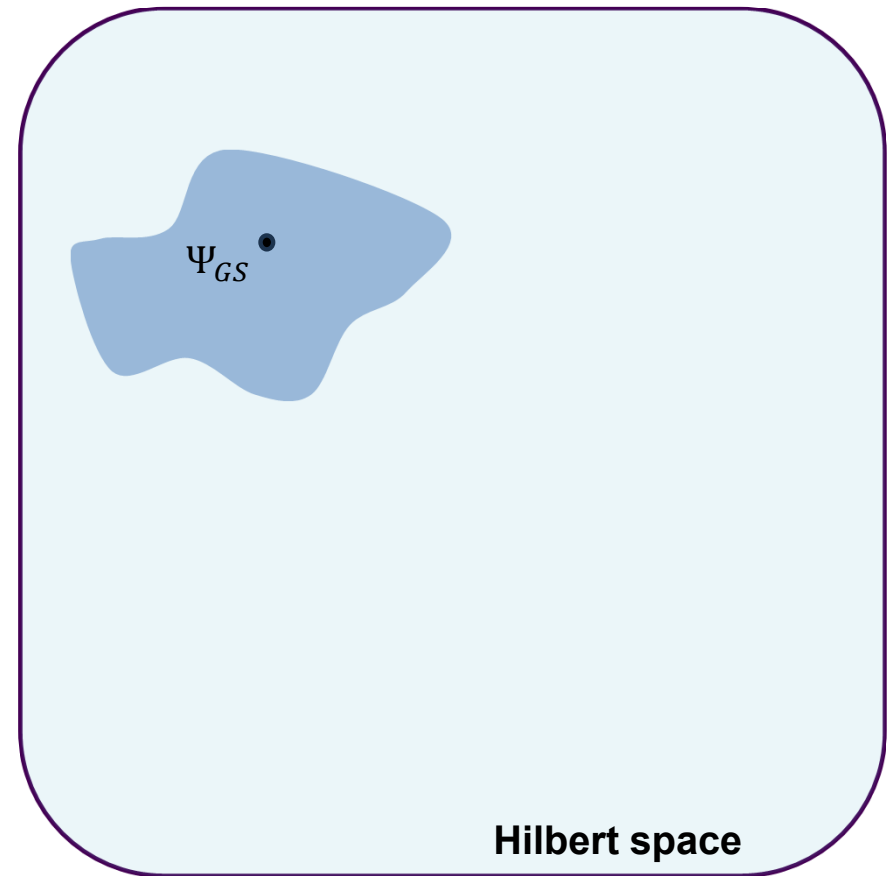
Variational wave-functions

Find the ground state

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$

Parametrize $|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S)|S\rangle$$



Variational wave-functions

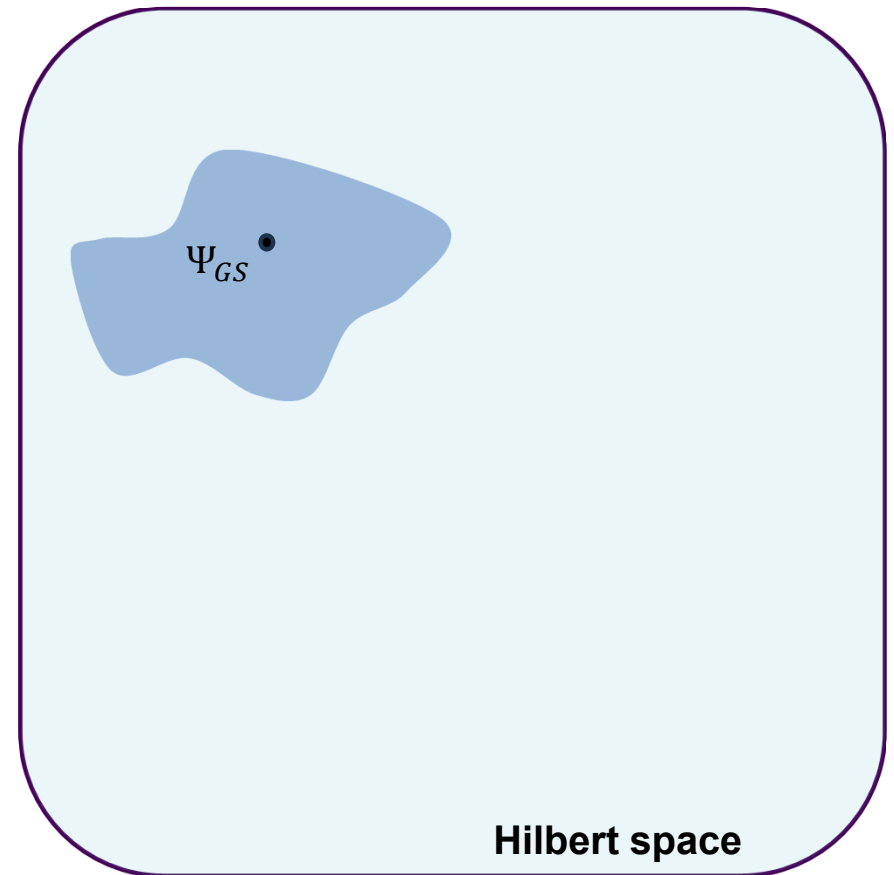
Find the ground state

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$

Parametrize $|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S)|S\rangle$$

Examples:



Variational wave-functions

Find the ground state

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$

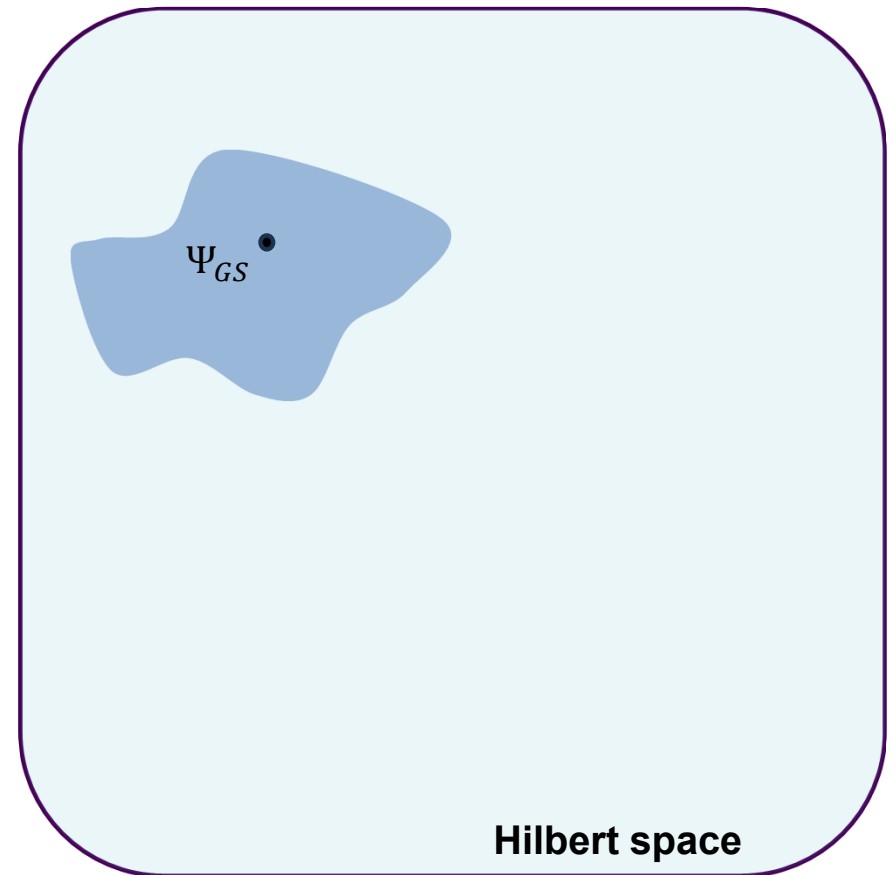
Parametrize $|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S)|S\rangle$$

Examples:

1) Product (mean-field) ansatz

$$\Psi_{\theta}(s_1, \dots, s_N) = \Psi_{\theta_1}(s_1) \cdot \Psi_{\theta_2}(s_2) \cdot \dots \cdot \Psi_{\theta_N}(s_N)$$



Variational wave-functions

Find the ground state

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$

Parametrize $|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S)|S\rangle$$

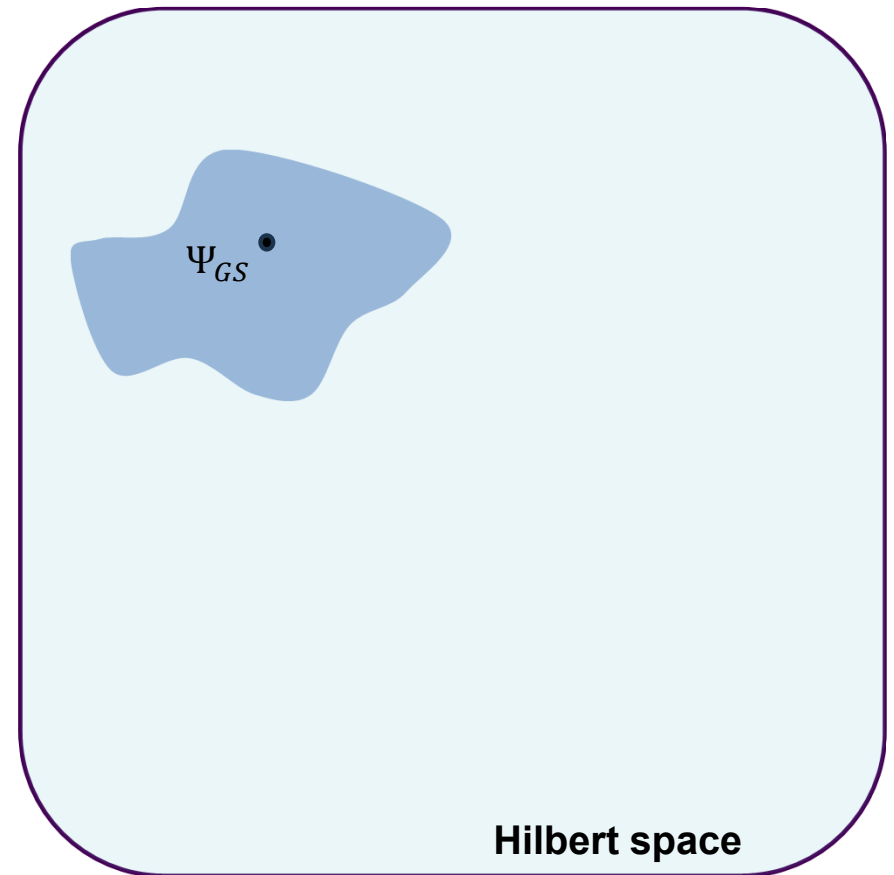
Examples:

1) Product (mean-field) ansatz

$$\Psi_{\theta}(s_1, \dots, s_N) = \Psi_{\theta_1}(s_1) \cdot \Psi_{\theta_2}(s_2) \cdot \dots \cdot \Psi_{\theta_N}(s_N)$$



E.g. can not represent $|\Psi\rangle_{GHZ} = |0\rangle^{\otimes N} + |1\rangle^{\otimes N}$
(product state is not entangled)



Variational wave-functions

Find the ground state

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$

Parametrize $|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S)|S\rangle$$

Examples:

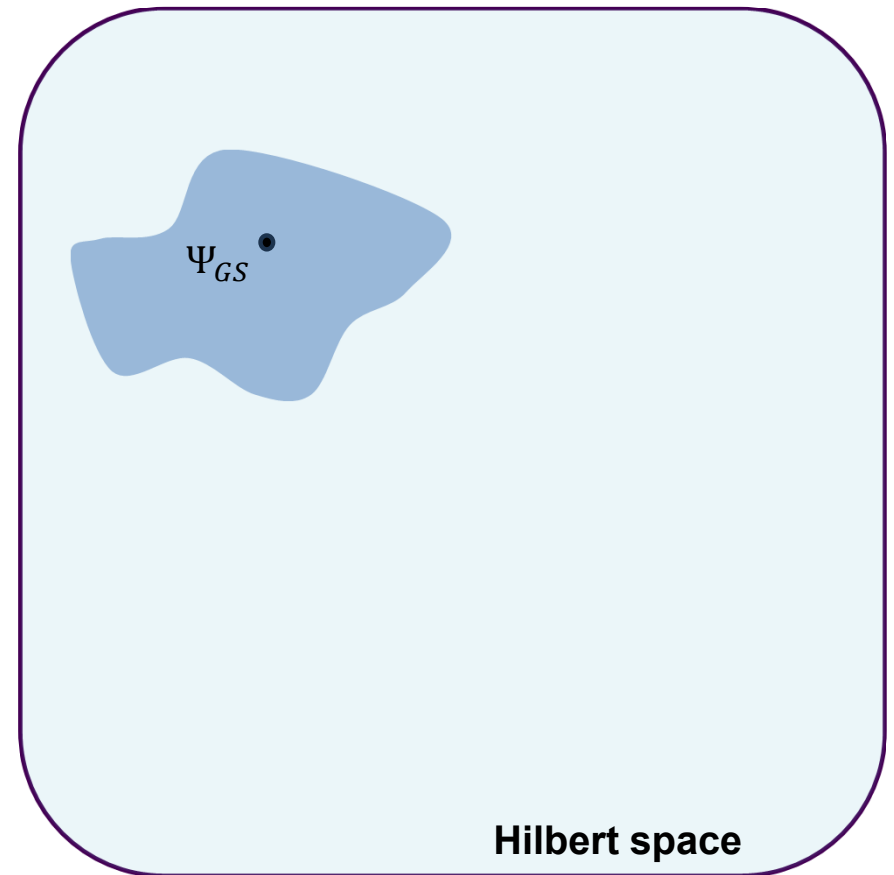
1) Product (mean-field) ansatz

$$\Psi_{\theta}(s_1, \dots, s_N) = \Psi_{\theta_1}(s_1) \cdot \Psi_{\theta_2}(s_2) \cdot \dots \cdot \Psi_{\theta_N}(s_N)$$

2) Matrix product states

$$\Psi_{\theta}(s_1, \dots, s_N) = \text{Tr}[\mathbf{A}_1^{s_1} \mathbf{A}_2^{s_2} \dots \mathbf{A}_N^{s_N}]$$

Area-law entanglement



Variational wave-functions

Find the ground state

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle$$

Parametrize $|\Psi\rangle$

$$|\Psi\rangle = \sum_S \Psi(S)|S\rangle$$

Examples:

1) Product (mean-field) ansatz

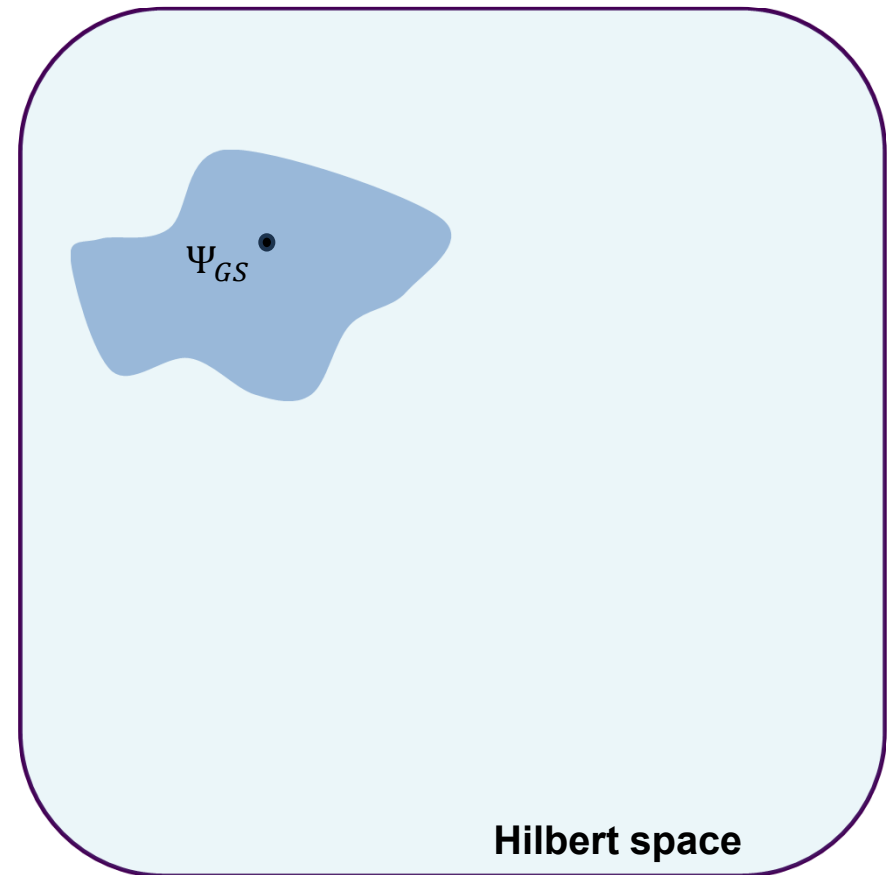
$$\Psi_{\theta}(s_1, \dots, s_N) = \Psi_{\theta_1}(s_1) \cdot \Psi_{\theta_2}(s_2) \cdot \dots \cdot \Psi_{\theta_N}(s_N)$$

2) Matrix product states

$$\Psi_{\theta}(s_1, \dots, s_N) = \text{Tr}[A_1^{s_1} A_2^{s_2} \dots A_N^{s_N}]$$

3) Neural-network quantum states

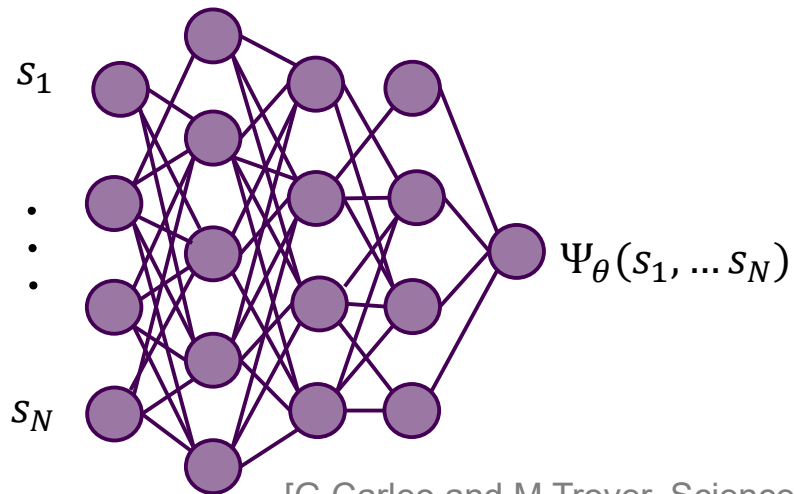
$$\Psi_{\theta}(s_1, \dots, s_N) = \text{NN}_{\theta}(s_1, \dots, s_N)$$



Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \underbrace{\Psi_\theta(s_1, \dots, s_N)}_{\text{Neural-net output}} |s_1, \dots, s_N\rangle$$

Neural-net output



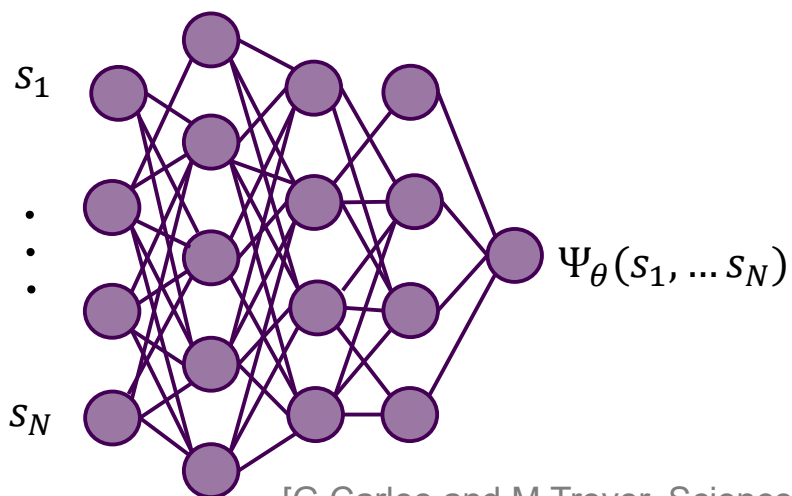
[G Carleo and M Troyer, Science (2017)]

[J Carrasquilla and R Melko, Nature Physics 13 (2017) - SI]

Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \underbrace{\Psi_\theta(s_1, \dots, s_N)}_{\text{Neural-net output}} |s_1, \dots, s_N\rangle$$

Neural-net output



[G Carleo and M Troyer, Science (2017)]

[J Carrasquilla and R Melko, Nature Physics 13 (2017) - SI]

What is a neural network?

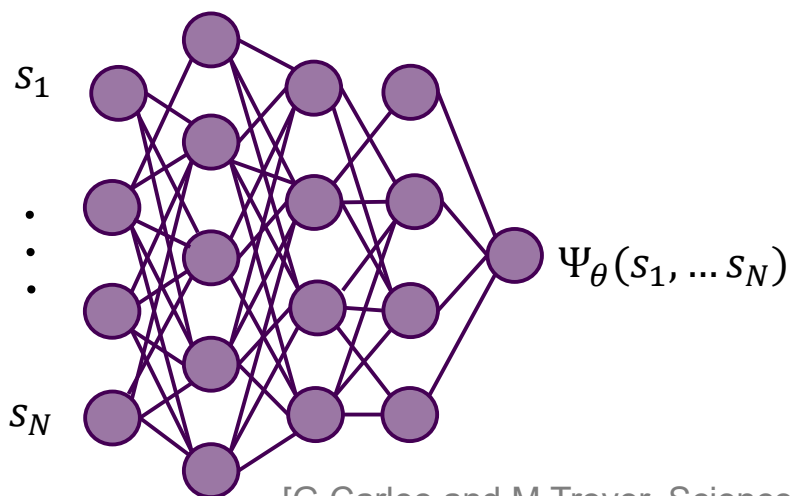
Output of a neuron =
nonlinear function of
weighted sum of inputs

The diagram shows three input nodes labeled y_1, y_2, y_3 on the left. Lines connect them to a central output node labeled $f(z)$. The connections are labeled with weights w_1, w_2, w_3 . The output node is labeled "output value". The weights are collectively labeled "weights".

Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \underbrace{\Psi_\theta(s_1, \dots, s_N)}_{\text{Neural-net output}} |s_1, \dots, s_N\rangle$$

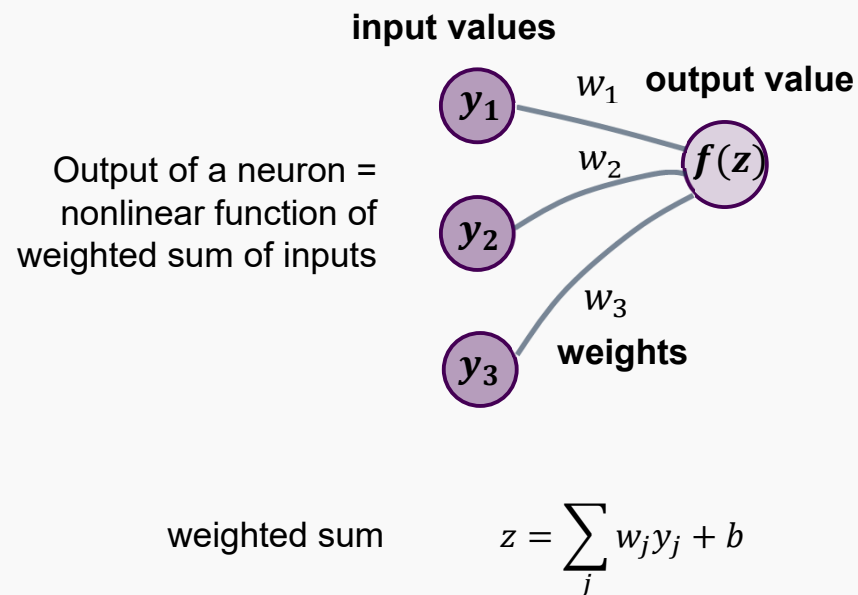
Neural-net output



[G Carleo and M Troyer, Science (2017)]

[J Carrasquilla and R Melko, Nature Physics 13 (2017) - SI]

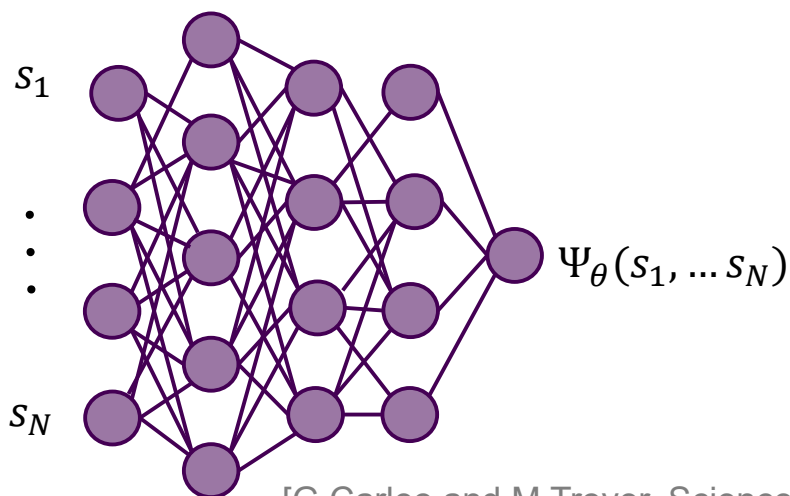
What is a neural network?



Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \underbrace{\Psi_\theta(s_1, \dots, s_N)}_{\text{Neural-net output}} |s_1, \dots, s_N\rangle$$

Neural-net output



[G Carleo and M Troyer, Science (2017)]

[J Carrasquilla and R Melko, Nature Physics 13 (2017) - SI]

What is a neural network?

Output of a neuron = nonlinear function of weighted sum of inputs

weighted sum

nonlinear "activation function"

input values

output value

weights

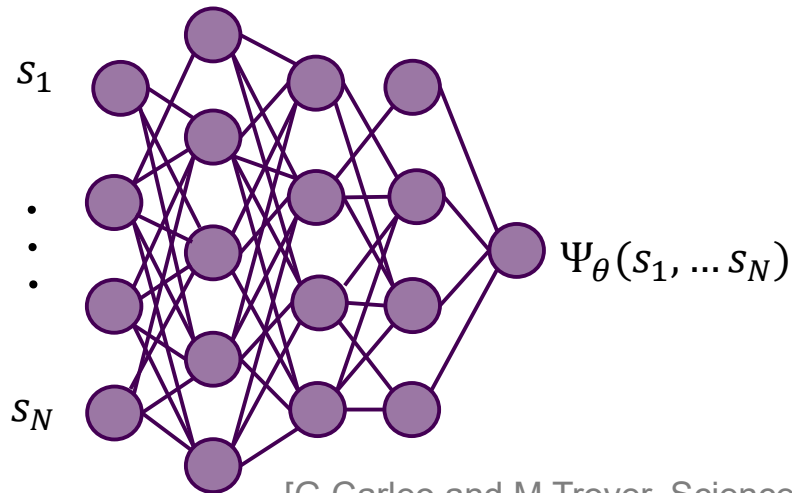
$$z = \sum_j w_j y_j + b$$

(rectified linear unit)

Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \underbrace{\Psi_\theta(s_1, \dots, s_N)}_{\text{Neural-net output}} |s_1, \dots, s_N\rangle$$

Neural-net output



[G Carleo and M Troyer, Science (2017)]

[J Carrasquilla and R Melko, Nature Physics 13
(2017) - SI]

Universal approximation theorem

single-layer neural network:

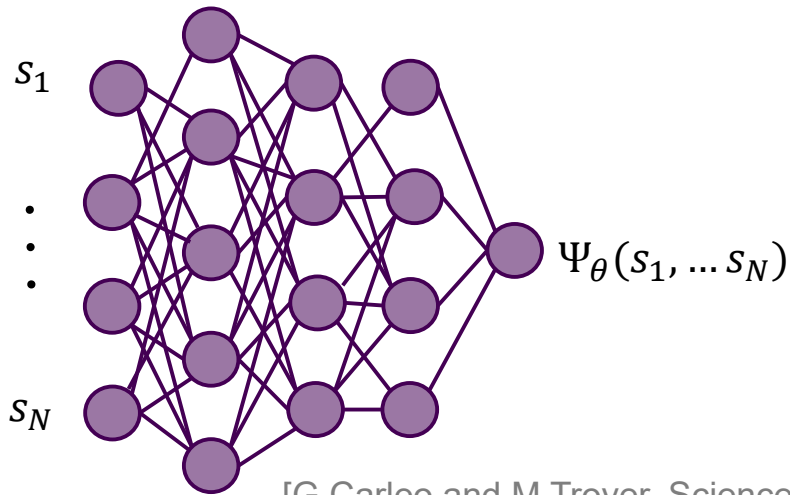
$$\text{Width} \rightarrow \infty \quad \Rightarrow \quad \|f^* - f_W\| \sim \text{width}^{-1}$$

[Cybenko, MCSS 2, 303 (1989)]

[Leshno et al., NN 6 861 (1993)]

Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \underbrace{\Psi_\theta(s_1, \dots, s_N)}_{\text{Neural-net output}} |s_1, \dots, s_N\rangle$$



[G Carleo and M Troyer, Science (2017)]

[J Carrasquilla and R Melko, Nature Physics 13 (2017) - SI]

Universal approximation theorem

single-layer neural network:

$$\text{Width} \rightarrow \infty \quad \Rightarrow \quad \|f^* - f_W\| \sim \text{width}^{-1}$$

[Cybenko, MCSS 2, 303 (1989)]

[Leshno et al., NN 6 861 (1993)]

Deep neural networks: $\|f^* - f_W\| \sim \exp[-\text{depth}]?$

Proven in some exotic cases

[Z. Lu et al, NIPS 30, 6231 (2017)]

Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \Psi_\theta(s_1, \dots, s_N) |s_1, \dots, s_N\rangle$$

Representability

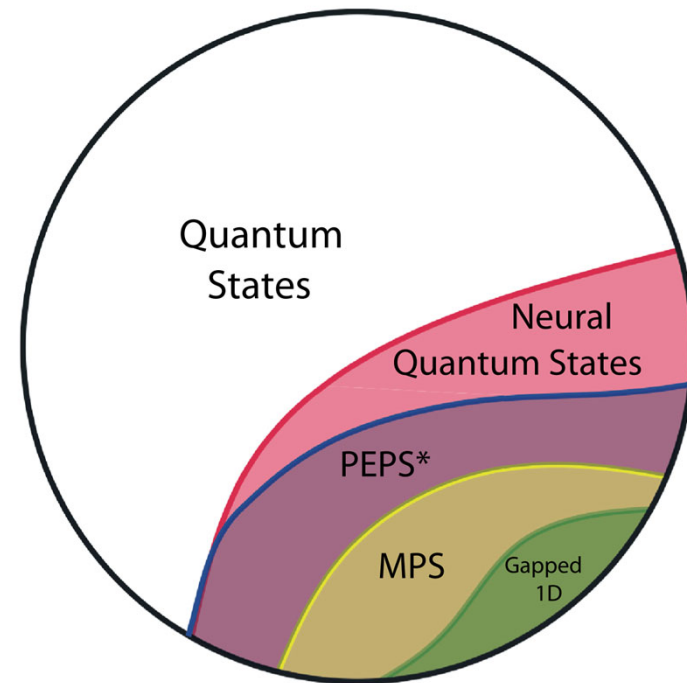
- Always valid quantum state (unnormalized)

Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \Psi_\theta(s_1, \dots, s_N) |s_1, \dots, s_N\rangle$$

Representability

- Always valid quantum state (unnormalized)
- Encode MPS and approximate PEPS with poly resources (BUT optimization harder + no tensor contraction) [Sharir, 2022]



[Sharir et al., PRB 106 (2022)]

[Carleo and Troyer, Science 355 (2017)]

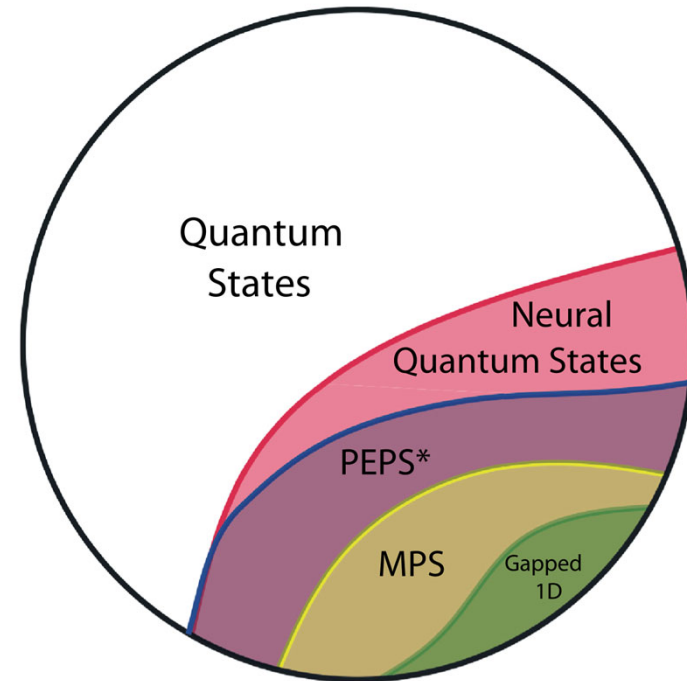
[Carleo et al, MRP RMP 91(2019)]

Neural-network quantum states

$$|\Psi\rangle = \sum_{s_1, \dots, s_N} \Psi_\theta(s_1, \dots, s_N) |s_1, \dots, s_N\rangle$$

Representability

- Always valid quantum state (unnormalized)
- Encode MPS and approximate PEPS with poly resources (BUT optimization harder + no tensor contraction) [Sharir, 2022]
- Volume-law in principle possible



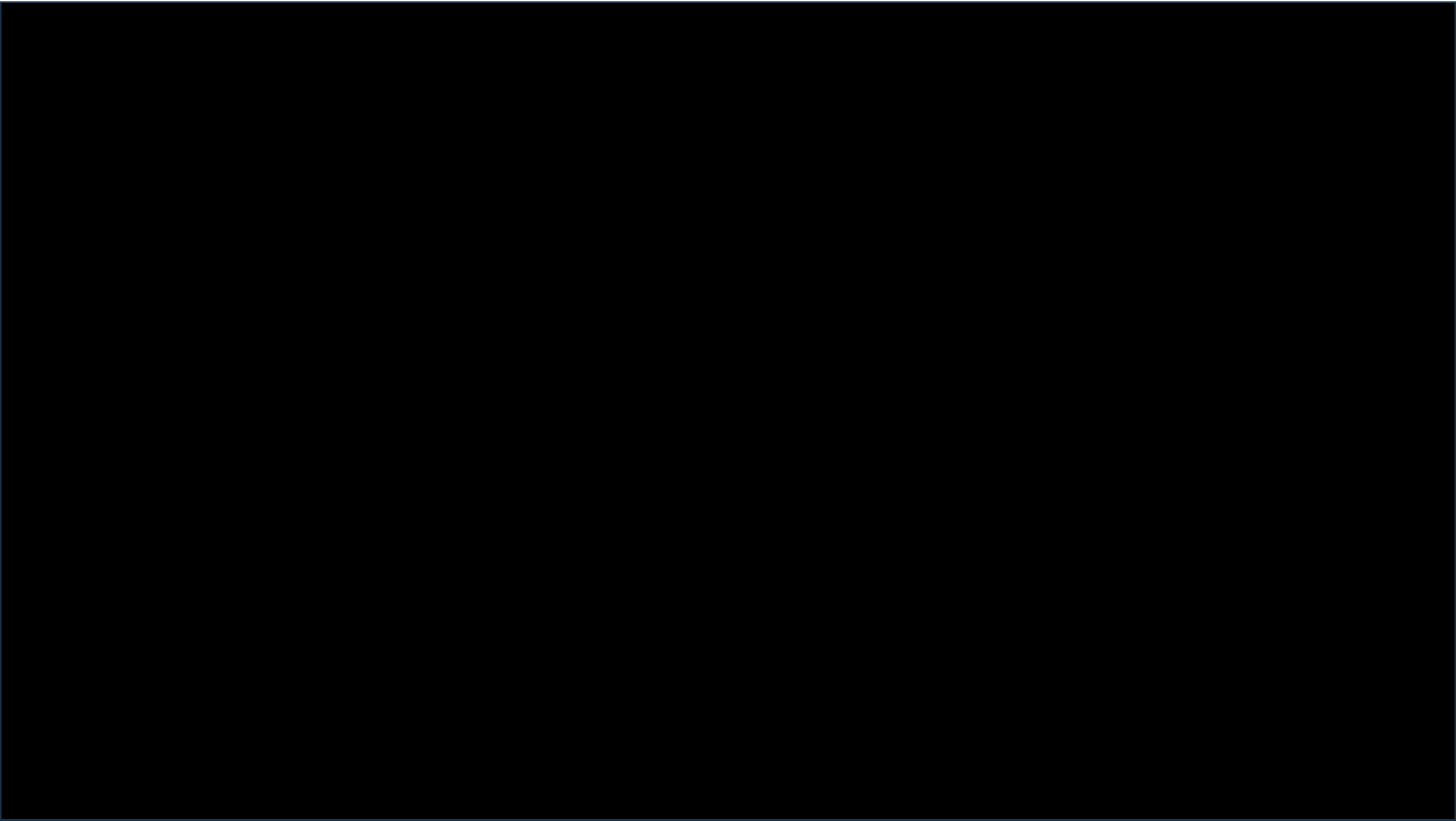
[Sharir et al., PRB 106 (2022)]

[Carleo and Troyer, Science 355 (2017)]

[Carleo et al, MRP RMP 91(2019)]

Part II

Variational Monte Carlo



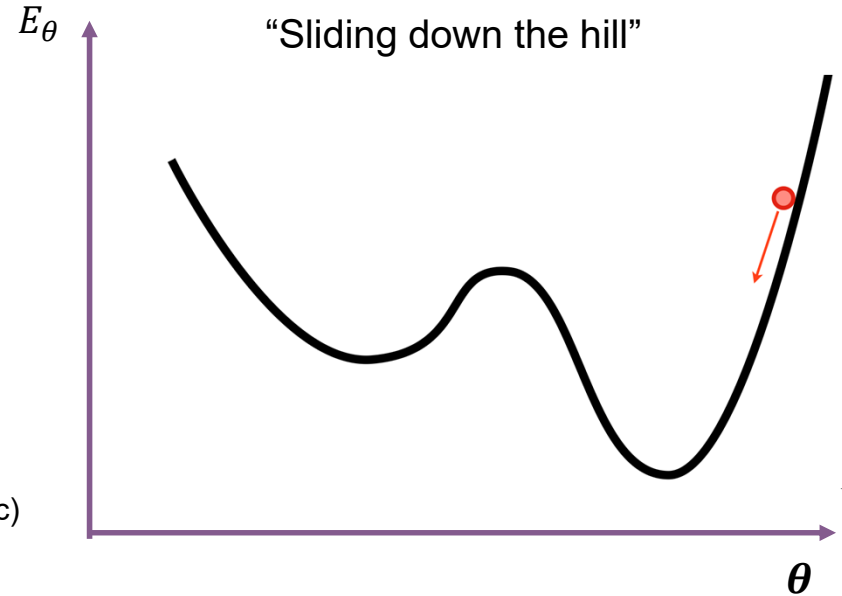
Stochastic optimization: ML perspective

Variational principle $E_\theta := \frac{\langle \Psi_\theta | H | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} \geq E_G$

(Stochastic) Gradient descent $\theta_{new} = \theta - \eta \nabla_\theta E_\theta$

with $\nabla_{\theta_k} E_\theta = \nabla_{\theta_k} \frac{\langle \Psi_\theta | H | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} = 2 \text{Re} \{ \underbrace{\langle O_k^\dagger H \rangle - \langle O_k^\dagger \rangle \langle H \rangle}_{\text{Estimated via sampling (stochastic)}} \}$

$O_k(S) = \frac{\partial \ln \Psi_\theta}{\partial \theta_k}(S)$



[Park and Kastoryano, PRR 2(2), 2020]

Not stable for NQS: Tends to oscillate between deep local minima

Improve by taking into account Riemannian metric

Stochastic optimization: ML perspective

Variational principle $E_\theta := \frac{\langle \Psi_\theta | H | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} \geq E_G$

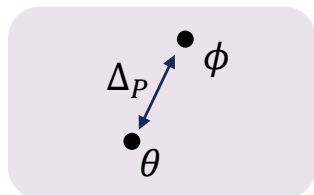
Natural gradient descent

$$\theta_{new} = \theta - \eta g^{-1} \nabla_\theta E_\theta$$

Infinitesimal distance

$$||d\theta||^2 = \sum_{\alpha\beta} \underbrace{g_{\alpha\beta}(\theta)}_{\text{Metric tensor}} d\theta_\alpha d\theta_\beta \quad \Rightarrow \quad \nabla_\theta \rightarrow \tilde{\nabla}_\theta = g^{-1} \nabla_\theta$$

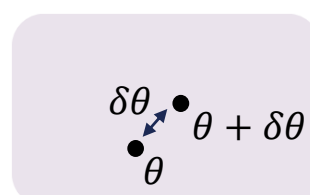
Parameter space



Euclidean distance

$$\Delta_P = |\vec{\theta} - \vec{\phi}|$$

Parameter space



$$ds^2 = \sum_{\alpha} d\theta_{\alpha}^2 = \sum_{\alpha,\beta} \delta_{\alpha\beta} d\theta_{\alpha} d\theta_{\beta}$$

→ "vanilla" gradient descent!

Stochastic optimization: ML perspective

Variational principle $E_\theta := \frac{\langle \Psi_\theta | H | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} \geq E_G$

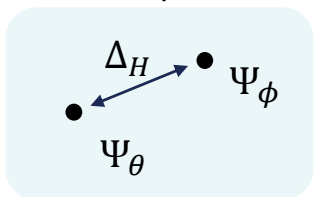
Natural gradient descent

$$\theta_{new} = \theta - \eta g^{-1} \nabla_\theta E_\theta$$

Infinitesimal distance

$$||d\theta||^2 = \sum_{\alpha\beta} \underbrace{g_{\alpha\beta}(\theta)}_{\text{Metric tensor}} d\theta_\alpha d\theta_\beta \quad \Rightarrow \quad \nabla_\theta \rightarrow \tilde{\nabla}_\theta = g^{-1} \nabla_\theta$$

Hilbert space

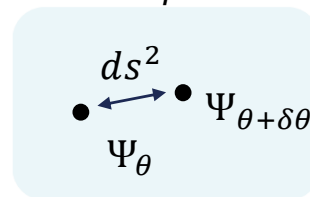


Fubini-Study distance

$$\Delta_H = \arccos |F_{\theta\phi}|,$$

$$F_{\theta\phi} = \langle \Psi_\theta | \Psi_\phi \rangle$$

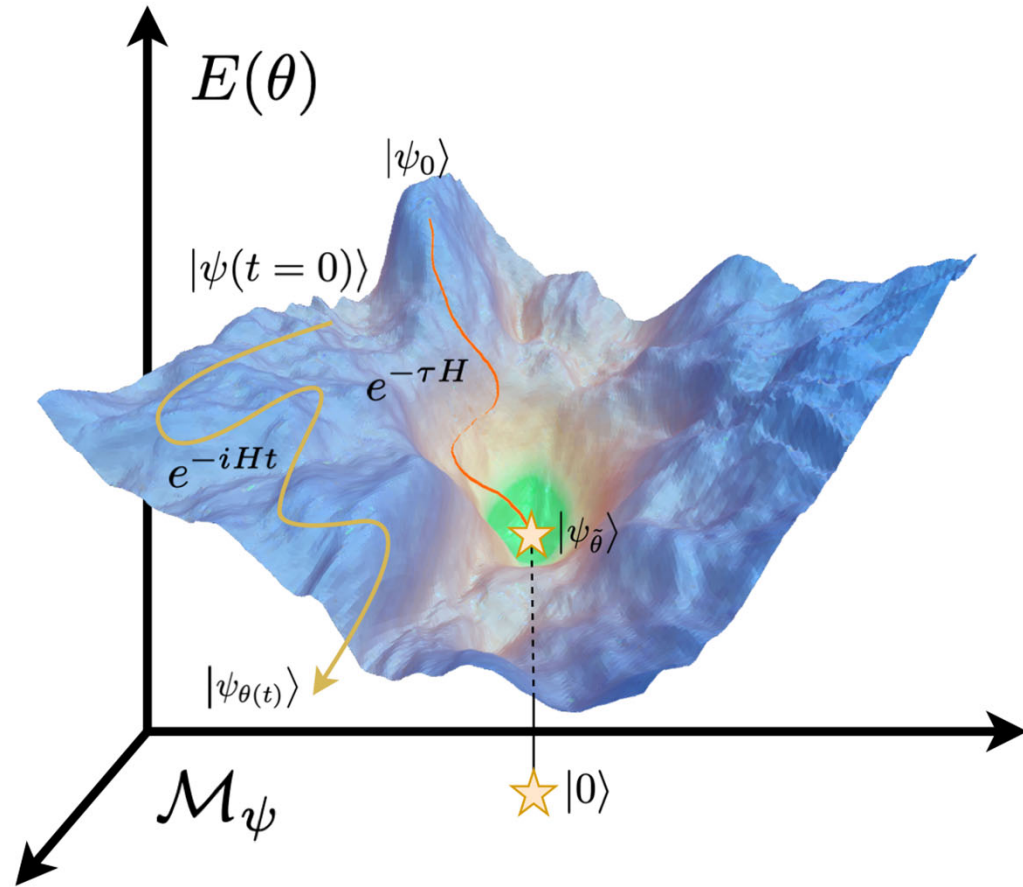
Hilbert space



$$ds^2 = \sum_{\alpha,\beta} S_{\alpha\beta} d\theta_\alpha d\theta_\beta$$

$$S_{\alpha\beta} = \left\langle \frac{\partial \psi_\theta}{\partial \theta_\alpha} \middle| \frac{\partial \psi_\theta}{\partial \theta_\alpha} \right\rangle - \left\langle \frac{\partial \psi_\theta}{\partial \theta_\alpha} \middle| \psi_\theta \right\rangle \left\langle \psi_\theta \middle| \frac{\partial \psi_\theta}{\partial \theta_\beta} \right\rangle$$

→ "Stochastic Reconfiguration"



Symmetries

Set of symmetry operations $SG = \{g\}$:

$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

↑
character

Symmetry sector i

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Symmetry sector i

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

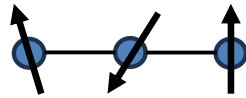
$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Symmetry sector i

Example:

3 spins + PBC



Symmetries

Set of symmetry operations $SG = \{g\}$:

$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

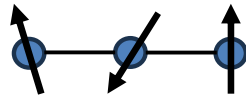
Symmetry sector i

Direct symmetrization:

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

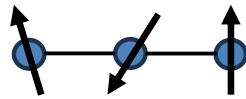
$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

Symmetry sector i

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Symmetry sector 0:

Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

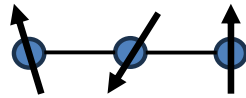
Symmetry sector i

Direct symmetrization:

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Symmetry sector 0:

$$\Psi_\theta^0(\mathbf{x}) = \tilde{\Psi}(\mathbf{x}) + \tilde{\Psi}(T^{-1}\mathbf{x}) + \tilde{\Psi}(T^{-2}\mathbf{x})$$

Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

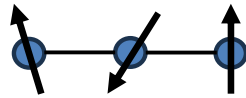
$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

Symmetry sector i

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetry sector 0:

$$\Psi_\theta^0(x) = \tilde{\Psi}(x) + \tilde{\Psi}(T^{-1}x) + \tilde{\Psi}(T^{-2}x)$$

$$\Psi_\theta^0(Tx) = \tilde{\Psi}(Tx) + \tilde{\Psi}(x) + \tilde{\Psi}(T^{-1}x) = \Psi_\theta^0(x)$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

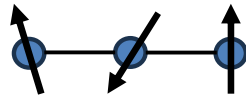
$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

Symmetry sector i

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetry sector 0:

$$\Psi_\theta^0(x) = \tilde{\Psi}(x) + \tilde{\Psi}(T^{-1}x) + \tilde{\Psi}(T^{-2}x)$$

$$\Psi_\theta^0(Tx) = \underbrace{\tilde{\Psi}(Tx) + \tilde{\Psi}(x) + \tilde{\Psi}(T^{-1}x)}_{\tilde{\Psi}(T^{-2}x)} = \Psi_\theta^0(x)$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

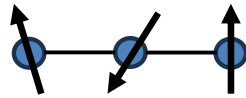
$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

Symmetry sector i

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetry sector 0:

$$\Psi_\theta^0(x) = \tilde{\Psi}(x) + \tilde{\Psi}(T^{-1}x) + \tilde{\Psi}(T^{-2}x)$$

$$\Psi_\theta^0(Tx) = \underbrace{\tilde{\Psi}(Tx) + \tilde{\Psi}(x) + \tilde{\Psi}(T^{-1}x)}_{\tilde{\Psi}(T^{-2}x)} = \Psi_\theta^0(x)$$

$$\Psi_\theta^0(T^2x) = \dots = \Psi_\theta^0(x)$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

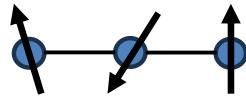
Symmetry sector i

Direct symmetrization:

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Symmetry sector 1:

Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

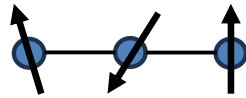
$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

Symmetry sector i

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetry sector 1:

$$\Psi_\theta^1(x) = \tilde{\Psi}(x) + \omega \tilde{\Psi}(T^{-1}x) + \omega^2 \tilde{\Psi}(T^{-2}x)$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

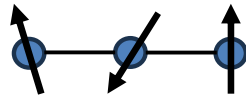
$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

Symmetry sector i

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Example:

3 spins + PBC



Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetry sector 1:

$$\Psi_\theta^1(x) = \tilde{\Psi}(x) + \omega \tilde{\Psi}(T^{-1}x) + \omega^2 \tilde{\Psi}(T^{-2}x)$$

$$\Psi_\theta^1(Tx) = \tilde{\Psi}(T^{-2}x) + \omega \tilde{\Psi}(x) + \omega^2 \tilde{\Psi}(T^{-1}x)$$

$$= \omega \Psi_\theta^0(x)$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

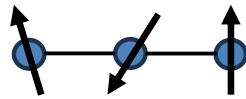
$$\Psi(g(x)) = \chi_g^i \Psi(x)$$

Symmetry sector i

$$\Psi_\theta^i(x) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}x)$$

Example:

3 spins + PBC



Translational symmetry:

	1	T	T^2
χ^0	1	1	1
χ^1	1	ω	ω^2
χ^2	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}}$$

Symmetry sector 1:

$$\Psi_\theta^1(x) = \tilde{\Psi}(x) + \omega \tilde{\Psi}(T^{-1}x) + \omega^2 \tilde{\Psi}(T^{-2}x)$$

$$\begin{aligned} \Psi_\theta^1(Tx) &= \tilde{\Psi}(T^{-2}x) + \omega \tilde{\Psi}(x) + \omega^2 \tilde{\Psi}(T^{-1}x) \\ &= \omega \Psi_\theta^0(x) \end{aligned}$$

$$\Psi_\theta^1(T^2x) = \dots = \omega^2 \Psi_\theta^1(x)$$

Symmetries

Set of symmetry operations $SG = \{g\}$:

Direct symmetrization:

$$\Psi(g(\mathbf{x})) = \chi_g^i \Psi(\mathbf{x})$$

Symmetry sector i

$$\Psi_\theta^i(\mathbf{x}) = \sum_g \chi_g^i \tilde{\Psi}(g^{-1}\mathbf{x})$$

Active area of research

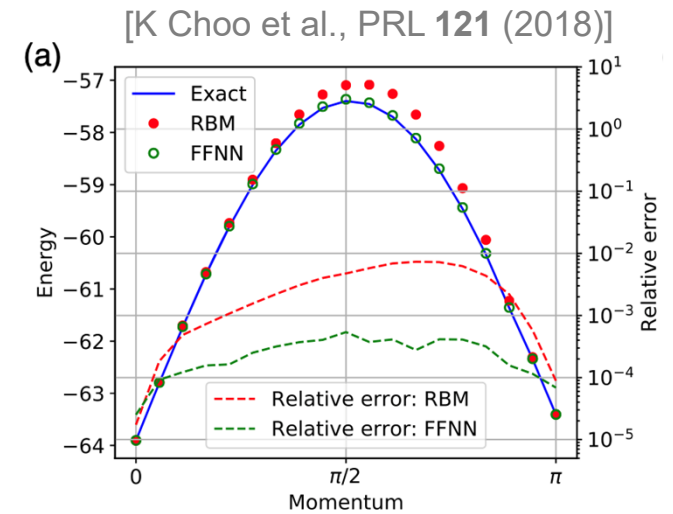
[T Vieijra et al., PRL 124 (2020)]

[T Vieijra and J Nys, PRB 104 (2021)]

[C Roth and A MacDonald, arXiv:2104 (2021)]

[M Reh et al., PRB 107 (2023)]

⋮



$$\text{AFM Heisenberg } H = \sum_{\langle i,j \rangle} S_i \cdot S_j$$

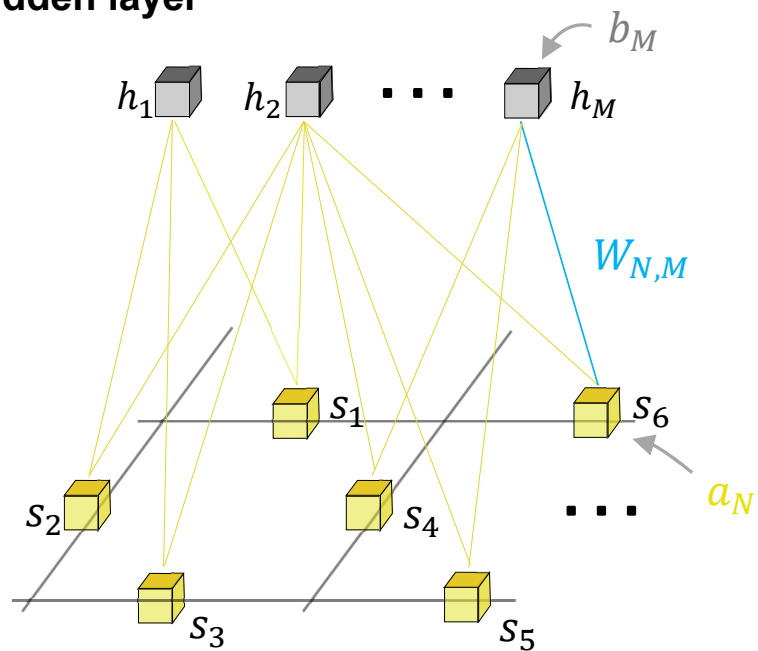
Part III

Applications: Architectures and benchmarks

Feed-forward architectures

Feed-forward architectures: Restricted Boltzmann Machine (RBM)

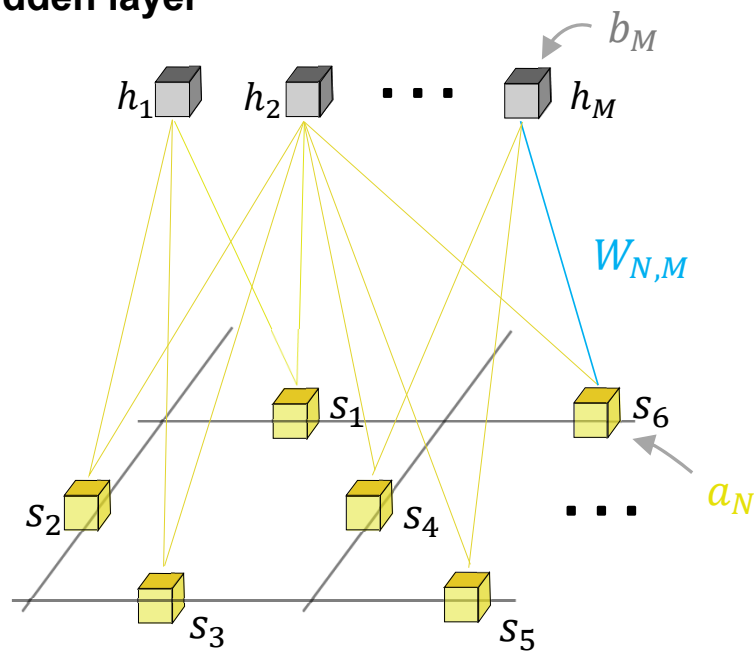
hidden layer



visible layer

Feed-forward architectures: Restricted Boltzmann Machine (RBM)

hidden layer

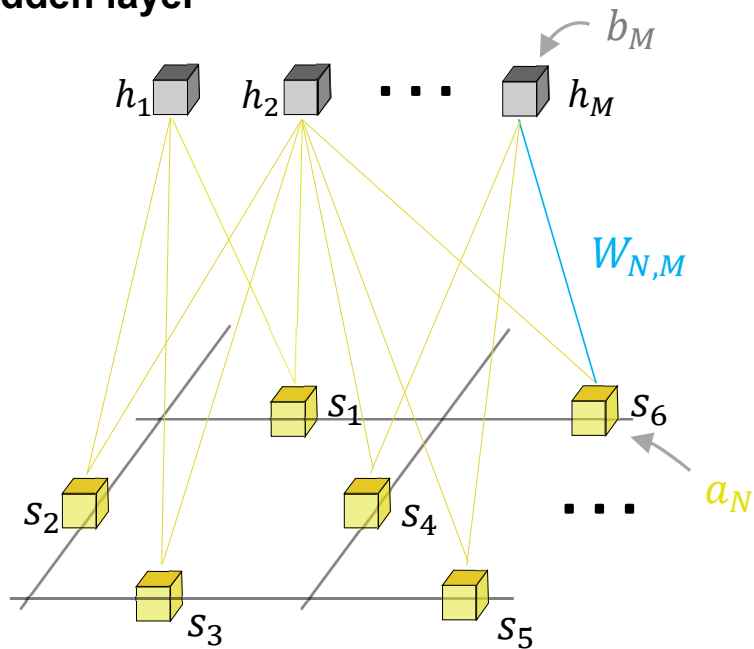


visible layer

- Variational state: $\Psi_{\theta}(S) = \sum_h \exp(E_{RBM})$,
 $E_{RBM} = \sum_k a_k s_k + \sum_j b_j h_j + \sum_{k,j} W_{k,j} s_k h_j$

Feed-forward architectures: Restricted Boltzmann Machine (RBM)

hidden layer



visible layer

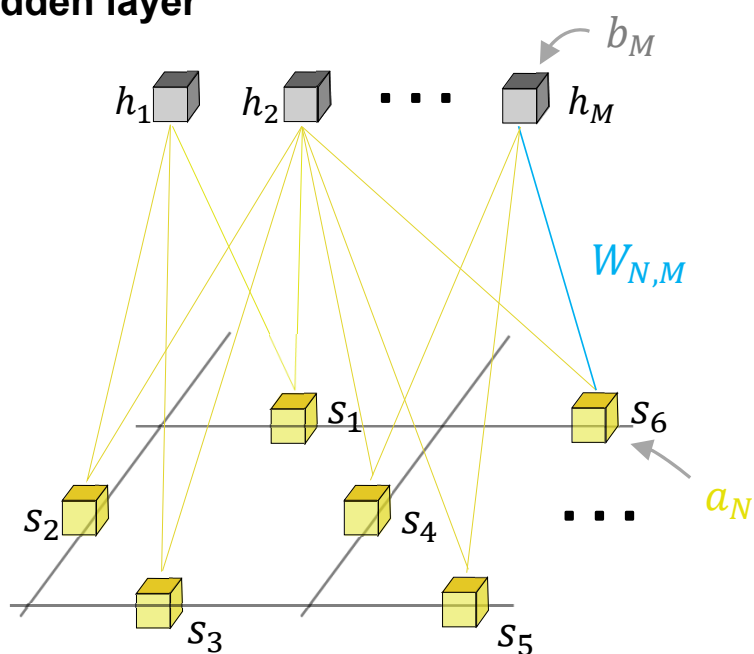
- Variational state: $\Psi_{\theta}(S) = \sum_h \exp(E_{RBM})$,
- $E_{RBM} = \sum_k a_k s_k + \sum_j b_j h_j + \sum_{k,j} W_{k,j} s_k h_j$

- "factorize":

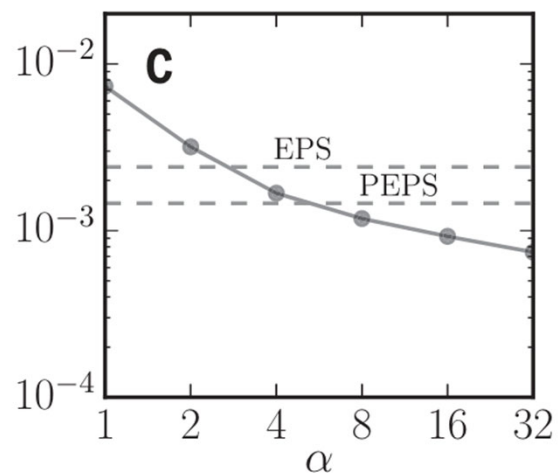
$$\Psi_{\theta}(S) = e^{\sum_k a_k s_k} \prod_j 2 \cosh \left(\sum_{k,j} W_{k,j} s_k h_j + b_j \right)$$

Feed-forward architectures: Restricted Boltzmann Machine (RBM)

hidden layer



visible layer



AF Heisenberg

$$H = \sum_{\langle i,j \rangle} S_i \cdot S_j$$

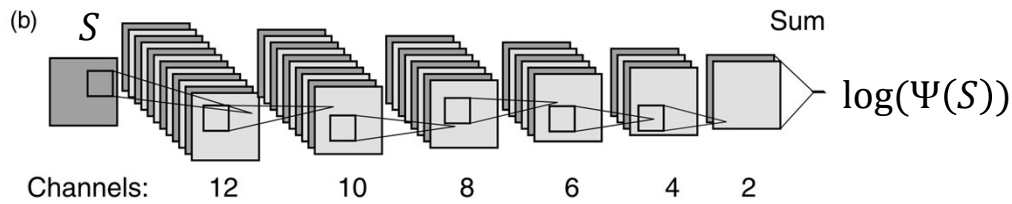
10 × 10 lattice

- Variational state: $\Psi_{\theta}(S) = \sum_h \exp(E_{RBM})$,
 $E_{RBM} = \sum_k a_k s_k + \sum_j b_j h_j + \sum_{k,j} W_{k,j} s_k h_j$

- "factorize":

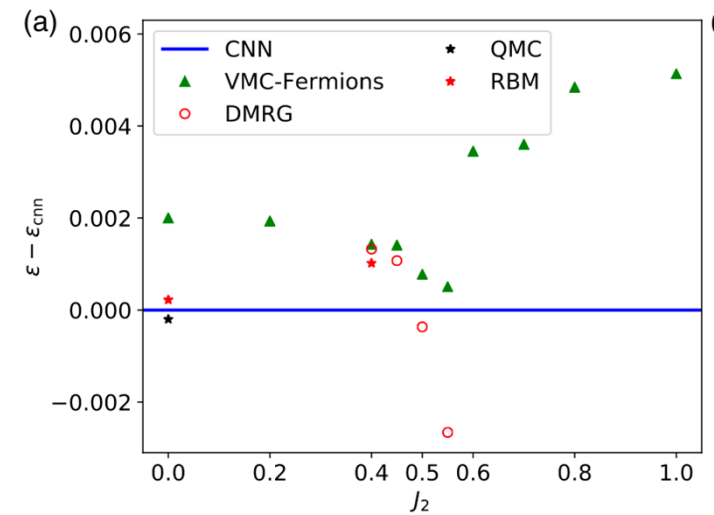
$$\Psi_{\theta}(S) = e^{\sum_k a_k s_k} \prod_j 2 \cosh \left(\sum_{k,j} W_{k,j} s_k h_j + b_j \right)$$

Feed-forward architectures: Convolutional NN (CNN)



Convolutional filters: Capitalize on translational symmetries

$$H = J_1 \sum_{\langle i,j \rangle} S_i \cdot S_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} S_i \cdot S_j$$



[K Choo et al., PRB 100 (2019)]

Recurrent neural networks: Autoregressive sampling

$$|\Psi\rangle = \sum_S \exp[i\phi(S)] \sqrt{P(S)} |S\rangle$$

\downarrow
 $P(S) = |\Psi(S)|^2$

$$P(S) = P(s_1)P(s_2|s_1) \cdot \dots \cdot P(s_N|s_{N-1}, \dots, s_2, s_1)$$

Autoregressive trick: Decompose probability distribution P into conditional probabilities

Recurrent neural networks: Autoregressive sampling

$$|\Psi\rangle = \sum_S \exp[i\phi(S)] \sqrt{P(S)} |S\rangle$$

$$P(S) = P(s_1)P(s_2|s_1) \cdot \dots \cdot P(s_N|s_{N-1}, \dots, s_2, s_1)$$

Autoregressive trick: Decompose probability distribution P into conditional probabilities

Sampling recipe:

1. Sample $s_i \sim p(s_i|s_{<i})$ and concatenate to $s_{<i}$
2. Use samples to define $p(s_{i+1}|s_{<i+1})$

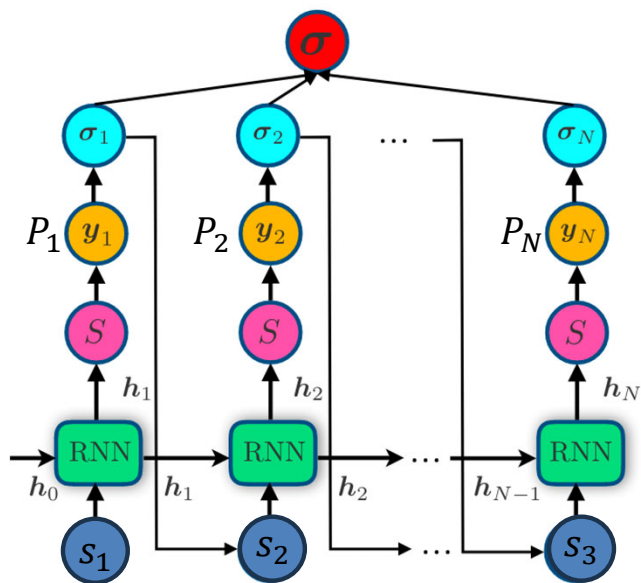
Obtain *uncorrelated* samples

Recurrent neural networks: Autoregressive sampling

$$|\Psi\rangle = \sum_S \exp[i\phi(S)] \sqrt{P(S)} |S\rangle$$

$$P(S) = P(s_1)P(s_2|s_1) \cdot \dots \cdot P(s_N|s_{N-1}, \dots, s_2, s_1)$$

Autoregressive trick: Decompose probability distribution P into conditional probabilities



Sampling recipe:

1. Sample $s_i \sim p(s_i | s_{<i})$ and concatenate to $s_{<i}$
2. Use samples to define $p(s_{i+1} | s_{<i+1})$

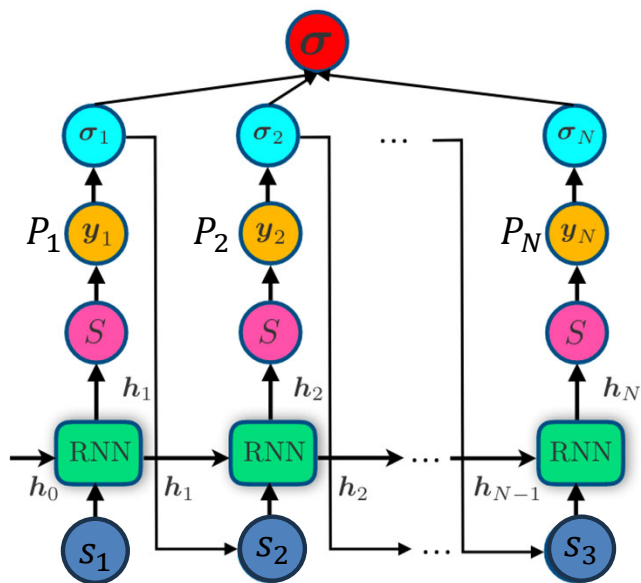
Obtain *uncorrelated* samples

Recurrent neural networks: Autoregressive sampling

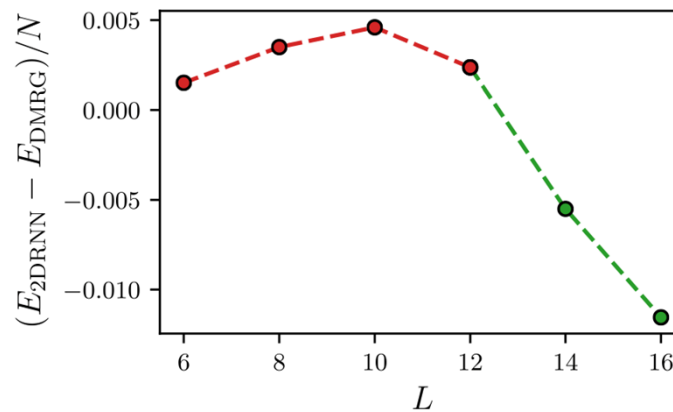
$$|\Psi\rangle = \sum_S \exp[i\phi(S)] \sqrt{P(S)} |S\rangle$$

$$P(S) = P(s_1)P(s_2|s_1) \cdot \dots \cdot P(s_N|s_{N-1}, \dots, s_2, s_1)$$

Autoregressive trick: Decompose probability distribution P into conditional probabilities



AFH Heisenberg on triangular lattice



[M Hibat-Allah et al., PRR 2 (2020)]

[M Hibat-Allah et al., arXiv:2207.14314 (2022)]

Recent advances: minSR

Stochastic reconfiguration:

$$\theta_{new} = \theta - \eta S^{-1} \nabla_{\theta} E_{\theta}$$

Quantum geometric tensor

$$S_{kk'} = \langle O_k^{\dagger} O_{k'} \rangle - \langle O_k^{\dagger} \rangle \langle O_{k'} \rangle \quad \text{with} \quad O_k(S) = \frac{\partial \ln \Psi_{\theta}}{\partial \theta_k}(S)$$

Recent advances: minSR

Stochastic reconfiguration:

$$\theta_{new} = \theta - \eta S^{-1} \nabla_{\theta} E_{\theta}$$

Quantum geometric tensor

$$S_{kk'} = \langle O_k^{\dagger} O_{k'} \rangle - \langle O_k^{\dagger} \rangle \langle O_{k'} \rangle \quad \text{with} \quad O_k(S) = \frac{\partial \ln \Psi_{\theta}}{\partial \theta_k}(S)$$



$[N_{parameters} \times$
 $N_{parameters}]$

Limit: GPU memory

Recent advances: minSR

Stochastic reconfiguration:

$$\theta_{new} = \theta - \eta S^{-1} \nabla_{\theta} E_{\theta}$$

Quantum geometric tensor

$$S_{kk'} = \langle O_k^{\dagger} O_{k'} \rangle - \langle O_k^{\dagger} \rangle \langle O_{k'} \rangle \quad \text{with} \quad O_k(S) = \frac{\partial \ln \Psi_{\theta}}{\partial \theta_k}(S)$$



$[N_{parameters} \times$
 $N_{parameters}]$

Limit: GPU memory

Linear algebra trick:
Instead invert $[N_{samples} \times N_{samples}]$ matrix
“minSR”

[A Chen and M Heyl, arXiv:2302.01941 (2023)]

[R Rende et al., arXiv:2310.05715 (2023)]

Recent advances: minSR

Stochastic reconfiguration:

$$\theta_{new} = \theta - \eta S^{-1} \nabla_{\theta} E_{\theta}$$

Quantum geometric tensor

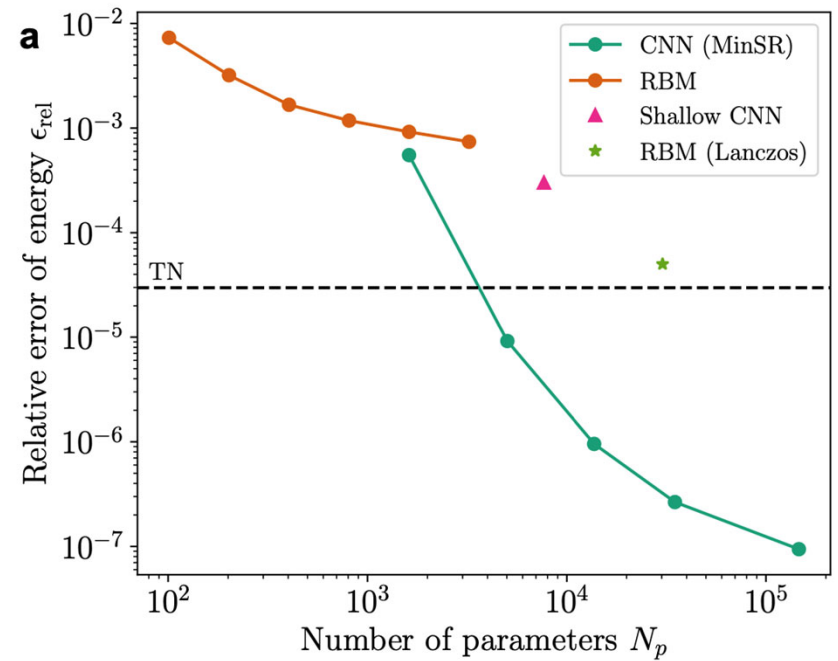
$$S_{kk'} = \langle O_k^{\dagger} O_{k'} \rangle - \langle O_k^{\dagger} \rangle \langle O_{k'} \rangle \quad \text{with} \quad O_k(S) = \frac{\partial \ln \Psi_{\theta}}{\partial \theta_k}(S)$$

↓
 $[N_{parameters} \times N_{parameters}]$

Limit: GPU memory

Linear algebra trick:
 Instead invert $[N_{samples} \times N_{samples}]$ matrix
 “minSR”

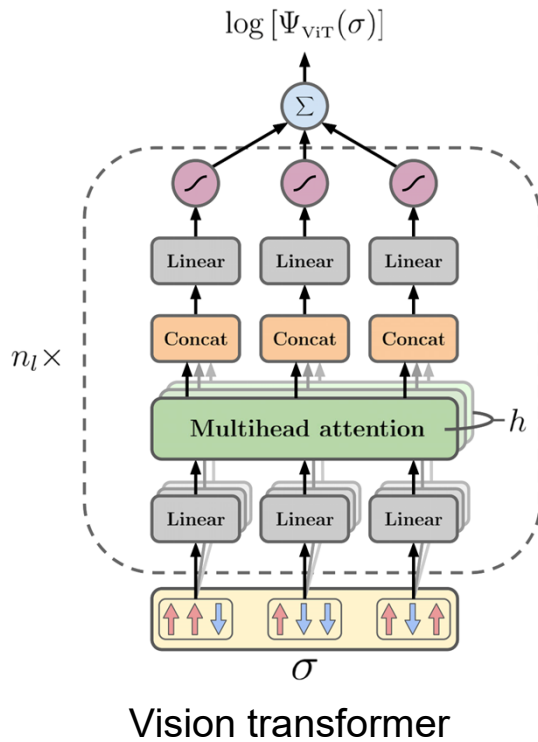
$$H = - \sum_{\langle i,j \rangle} S_i \cdot S_j$$



[A Chen and M Heyl, arXiv:2302.01941 (2023)]

[R Rende et al., arXiv:2310.05715 (2023)]

Benchmark energies



$$\log \Psi(S) = \text{RBM}(\mathcal{T}(S))$$

$$H = J_1 \sum_{\langle i,j \rangle} S_i \cdot S_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} S_i \cdot S_j$$

TABLE I. Ground-state energy on the 10×10 square lattice at $J_2/J_1 = 0.5$.

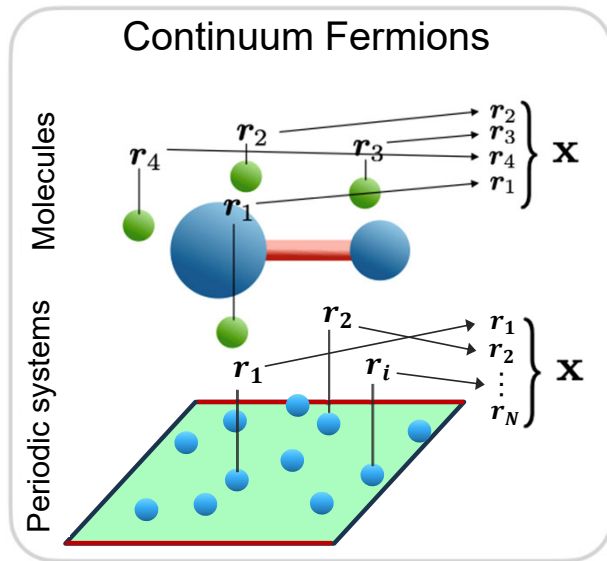
Energy per site	Wave function	# parameters	Marshall prior	Reference	Year
-0.48941(1)	NNQS	893994	Not available	[32]	2023
-0.494757(12)	CNN	Not available	No	[22]	2020
-0.4947359(1)	Shallow CNN	11009	Not available	[21]	2018
-0.49516(1)	Deep CNN	7676	Yes	[20]	2019
-0.495502(1)	PEPS + Deep CNN	3531	No	[33]	2021
-0.495530	DMRG	8192 SU(2) states	No	[31]	2014
-0.495627(6)	aCNN	6538	Yes	[34]	2023
-0.49575(3)	RBM-fermionic	2000	Yes	[15]	2019
-0.49586(4)	CNN	10952	Yes	[35]	2023
-0.4968(4)	RBM ($p = 1$)	Not available	Yes	[36]	2022
-0.49717(1)	Deep CNN	106529	Yes	[28]	2022
-0.497437(7)	GCNN	Not available	No	[27]	2021
-0.497468(1)	Deep CNN	421953	Yes	[30]	2022
-0.4975490(2)	VMC ($p = 2$)	5	Yes	[13]	2013
-0.497627(1)	Deep CNN	146320	Yes	[29]	2023
-0.497629(1)	RBM+PP	5200	Yes	[37]	2021
-0.497634(1)	Deep ViT	267720	No	Present work	2023

[R Rende et al., arXiv:2310.05715 (2023)]

Part IV

Fermionic systems

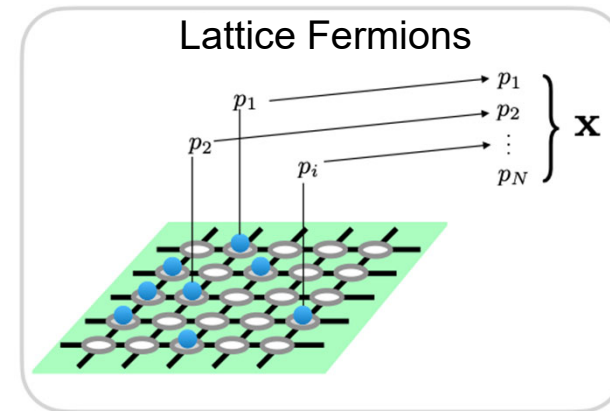
Fermions



Molecules, solids (full ab-initio),
electron gas, ...

[C Kenny et al., *Nature communications* 11.1 (2020)]

[J Robledo Moreno et al., *PNAS* 119 (32)]



Effective lattice models of materials
(Hubbard-like models)
Quantum simulators: Ultracold atoms

Fermions (continuum)

Fermions (continuum)

Antisymmetry

First quantization

$$\Psi(r_1, r_2, \dots, r_N) = -\Psi(r_2, r_1, \dots, r_N)$$

Second quantization

$$\{\hat{c}_\alpha^\dagger, \hat{c}_\beta^\dagger\} = \{\hat{c}_\alpha, \hat{c}_\beta\} = 0, \quad \{\hat{c}_\alpha, \hat{c}_\beta^\dagger\} = \delta_{\alpha\beta},$$



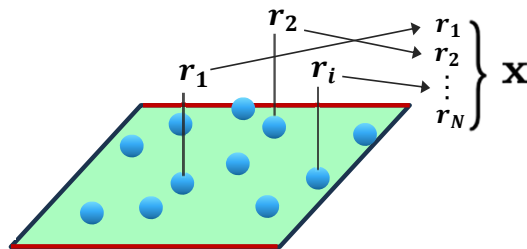
Creates single-particle state $|\phi_\alpha\rangle$,
e.g. $\phi_\alpha(\mathbf{r}_i) = e^{ik_\alpha \cdot \mathbf{r}_i}$

Fermions (continuum)

Antisymmetry

First quantization

$$\Psi(r_1, r_2, \dots, r_N) = -\Psi(r_2, r_1, \dots, r_N)$$



$\Psi(\mathbf{x})$ needs to be constructed such that antisymmetry is fulfilled

e.g. using a **Slater determinant**

Choice of basis

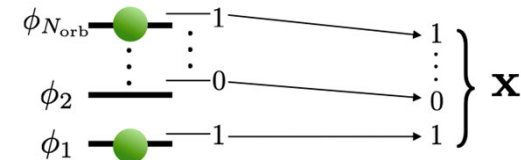
Second quantization

$$\{\hat{c}_\alpha^\dagger, \hat{c}_\beta^\dagger\} = \{\hat{c}_\alpha, \hat{c}_\beta\} = 0, \quad \{\hat{c}_\alpha, \hat{c}_\beta^\dagger\} = \delta_{\alpha\beta},$$



Creates single-particle state $|\phi_\alpha\rangle$,

e.g. $\phi_\alpha(\mathbf{r}_i) = e^{ik_\alpha \cdot \mathbf{r}_i}$



$\Psi(\mathbf{x})$ antisymmetric by construction of \mathbf{x}

Variational wave-functions for fermions (continuum)

Non-interacting wave-function (Hartree-Fock):

$$\Psi(r_1, \dots, r_N) = \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_\alpha(r_i)\}_{\alpha i})$$

Slater-Jastrow-Backflow wave-functions

Include correlations:

$$\Psi(r_1, \dots, r_N) = e^{-J(r_1, \dots, r_N)} \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_j(\tilde{r}_i)\}_{ij})$$



Jastrow factor

$$J(r_1, \dots, r_N) = \sum_{ij} J_{ij}(|r_i - r_j|)$$

Slater-Jastrow-Backflow wave-functions

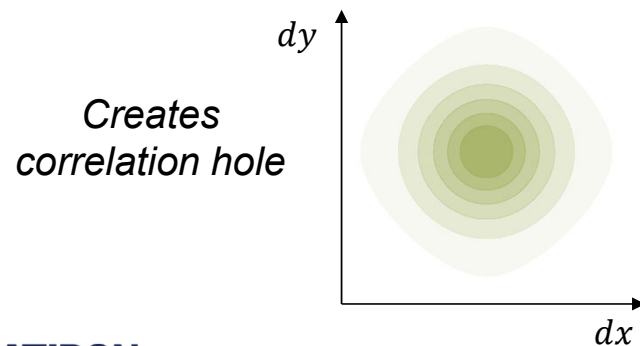
Include correlations:

$$\Psi(r_1, \dots, r_N) = e^{-J(r_1, \dots, r_N)} \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_j(\tilde{\mathbf{r}}_i)\}_{ij})$$



Jastrow factor

$$J(r_1, \dots, r_N) = \sum_{ij} J_{ij}(|\mathbf{r}_i - \mathbf{r}_j|)$$



Slater-Jastrow-Backflow wave-functions

Include correlations:

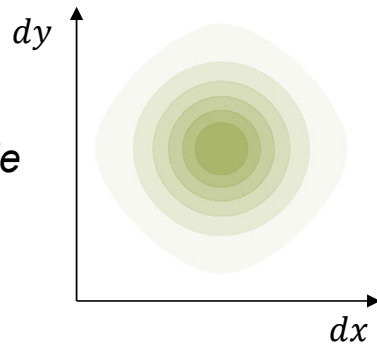
$$\Psi(r_1, \dots, r_N) = e^{-J(r_1, \dots, r_N)} \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_j(\tilde{r}_i)\}_{ij})$$



Jastrow factor

$$J(r_1, \dots, r_N) = \sum_{ij} J_{ij}(|r_i - r_j|)$$

Creates
correlation hole



Backflow transformation

$$\tilde{r}_i = r_i + \sum_j \xi_{\tau_i \tau_j}(|r_i - r_j|)(r_j - r_i)$$

Slater-Jastrow-Backflow wave-functions

Include correlations:

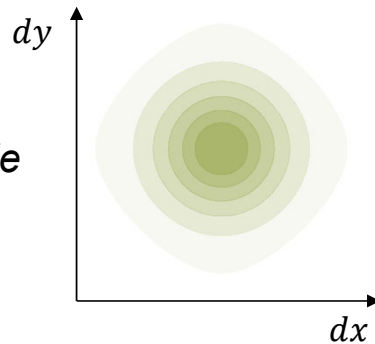
$$\Psi(r_1, \dots, r_N) = e^{-J(r_1, \dots, r_N)} \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_j(\tilde{r}_i)\}_{ij})$$



Jastrow factor

$$J(r_1, \dots, r_N) = \sum_{ij} J_{ij}(|r_i - r_j|)$$

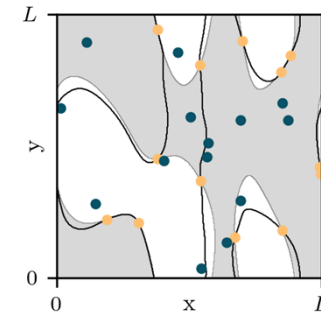
Creates correlation hole



Backflow transformation

$$\tilde{r}_i = r_i + \sum_j \xi_{\tau_i \tau_j} (|r_i - r_j|)(r_j - r_i)$$

Modifies nodal structure



Slater-Jastrow-Backflow wave-functions

Include correlations:

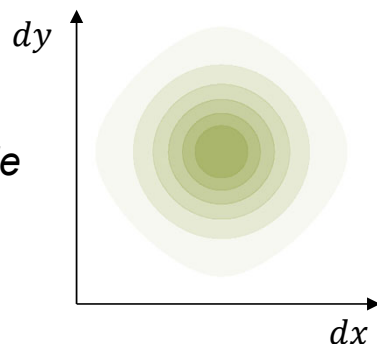
$$\Psi(r_1, \dots, r_N) = e^{-J(r_1, \dots, r_N)} \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_j(\tilde{\mathbf{r}}_i)\}_{ij})$$



Jastrow factor

$$J(r_1, \dots, r_N) = \sum_{ij} J_{ij}(|\mathbf{r}_i - \mathbf{r}_j|)$$

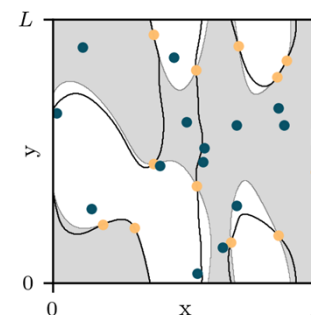
Creates correlation hole



Backflow transformation

$$\tilde{\mathbf{r}}_i = \mathbf{r}_i + \sum_j \xi_{\tau_i \tau_j}(|\mathbf{r}_i - \mathbf{r}_j|)(\mathbf{r}_j - \mathbf{r}_i)$$

*Modifies nodal structure
can in principle represent generic wave-functions*



Slater-Jastrow-Backflow wave-functions

Include correlations:

$$\Psi(r_1, \dots, r_N) = e^{-J(r_1, \dots, r_N)} \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_j(\tilde{r}_i)\}_{ij})$$

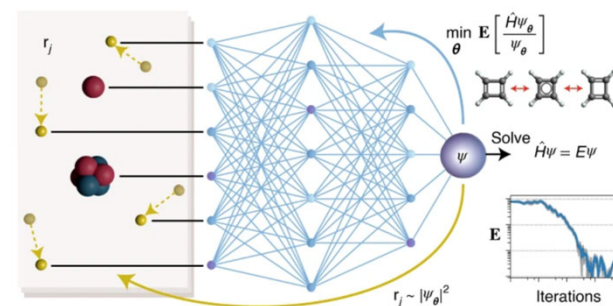


Backflow transformation

$$\tilde{r}_i = r_i + \underbrace{\delta r_i(r_1, \dots, r_N)}_{\text{NN}}$$

Parametrize with NN

[J Hermann et al.,
Nature Chemistry 12
(2020)]



Slater-Jastrow-Backflow wave-functions

Include correlations:

$$\Psi(r_1, \dots, r_N) = e^{-J(r_1, \dots, r_N)} \frac{1}{\sqrt{N!}} \text{Det}(\{\phi_j(\tilde{r}_i)\}_{ij})$$



Iterative backflow

[D Luo and B Clark, PRL 122 (2019)]

FermiNet

[D Pfau et al., PRR 2 (2020)]

[JS Spencer et al., arxiv:2011.07125 (2020)]

PauliNet

[J Hermann et al., Nature Chemistry 12 (2020)]

Graph NN

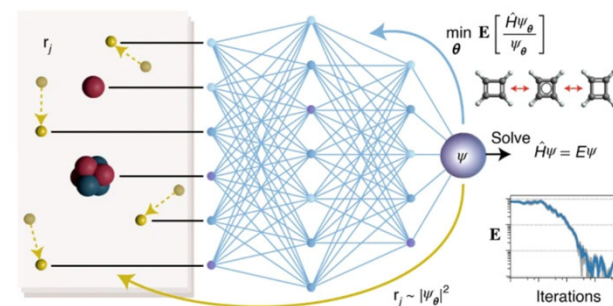
[Pescia et al., arXiv:2305.07240 (2023)]

Backflow transformation

$$\tilde{r}_i = r_i + \underbrace{\delta r_i(r_1, \dots, r_N)}_{\text{Parametrize with NN}}$$

Parametrize with NN

[J Hermann et al., Nature Chemistry 12 (2020)]

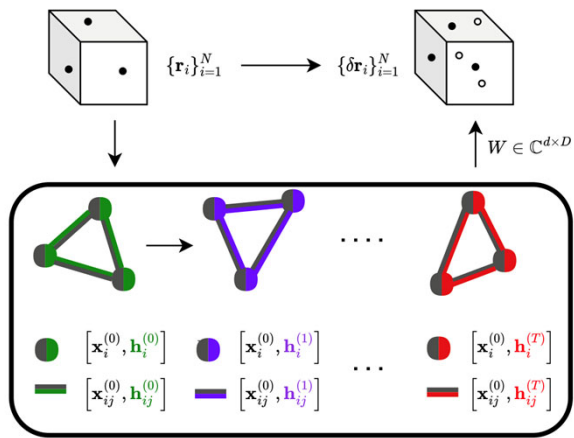


Example: 3D electron gas with message-passing NQS

Backflow transformation

$$\tilde{r}_i = r_i + \underbrace{\delta r_i(r_1, \dots, r_N)}$$

Parametrize with Message-passing Graph NN

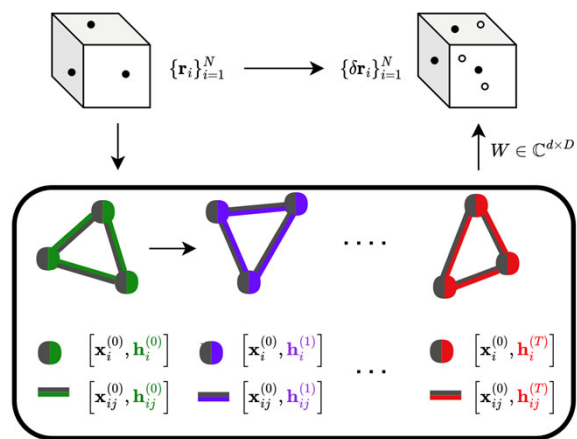


Example: 3D electron gas with message-passing NQS

Backflow transformation

$$\tilde{\mathbf{r}}_i = \mathbf{r}_i + \underbrace{\delta \mathbf{r}_i(\mathbf{r}_1, \dots, \mathbf{r}_N)}$$

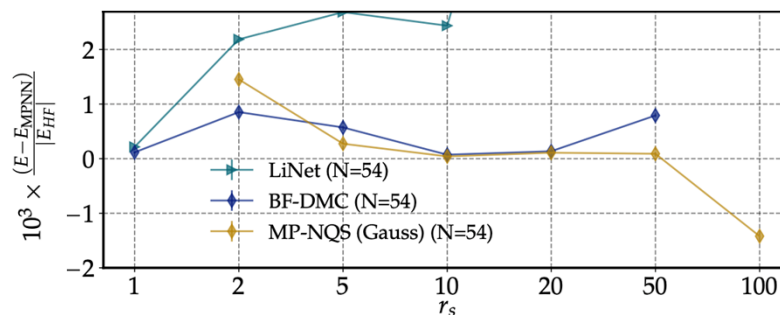
Parametrize with Message-passing Graph NN



3D interacting Electron gas

$$H = -\frac{1}{2r_s^2} \sum_i \nabla_i^2 + \frac{1}{r_s} \sum_{i < j} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|}$$

Wigner-Seitz radius $r_s = \sqrt[3]{3/(4\pi n)}$



[Pescia et al., arXiv:2305.07240 (2023)]

Outlook: Other applications

Real time evolution

[M Schmitt and M Heyl, PRL 125 (2020)]

[M Medvidovic and D Sels, PRX Quantum 4 (2023)]

[A Sinibaldi et al., arXiv2305.14294 (2023)]

Open systems

[A. Nagy and V. Savona, Phys. Rev. Lett. 122 (2019)]

[J. Carrasquilla et al., Nature Machine Intelligence 1 (2019)]

[F Vicentini et al., arXiv:2206.13488 (2022)]

Finite temperature

[N Irikura and Hiroki Saito, PRR 2 (2020)]

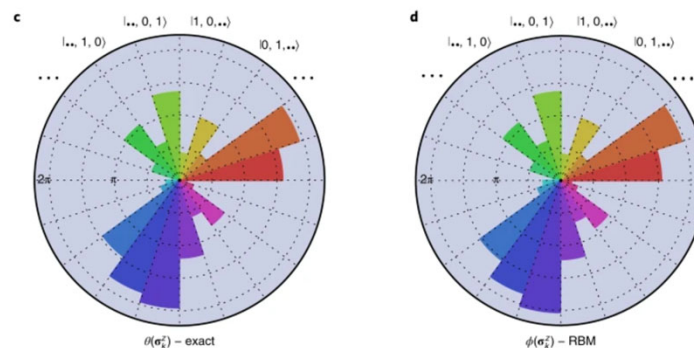
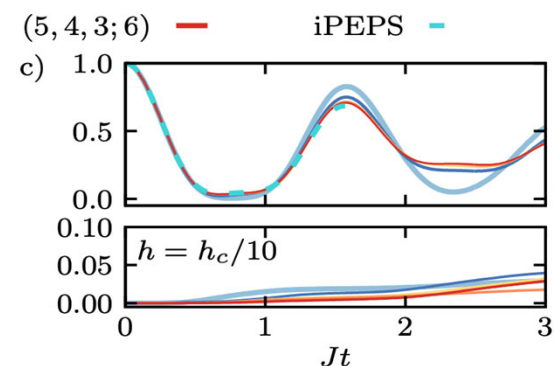
[Y Nomura et al., PRL 127 (2021)]

Quantum state reconstruction

[G Torlai et al., Nature physics 14 (2018)]

[J Carrasquilla et al., Nature Machine Intelligence 1.3 (2019)]

[S Czischek et al., PRB 105 (2022)]



Thank you.



