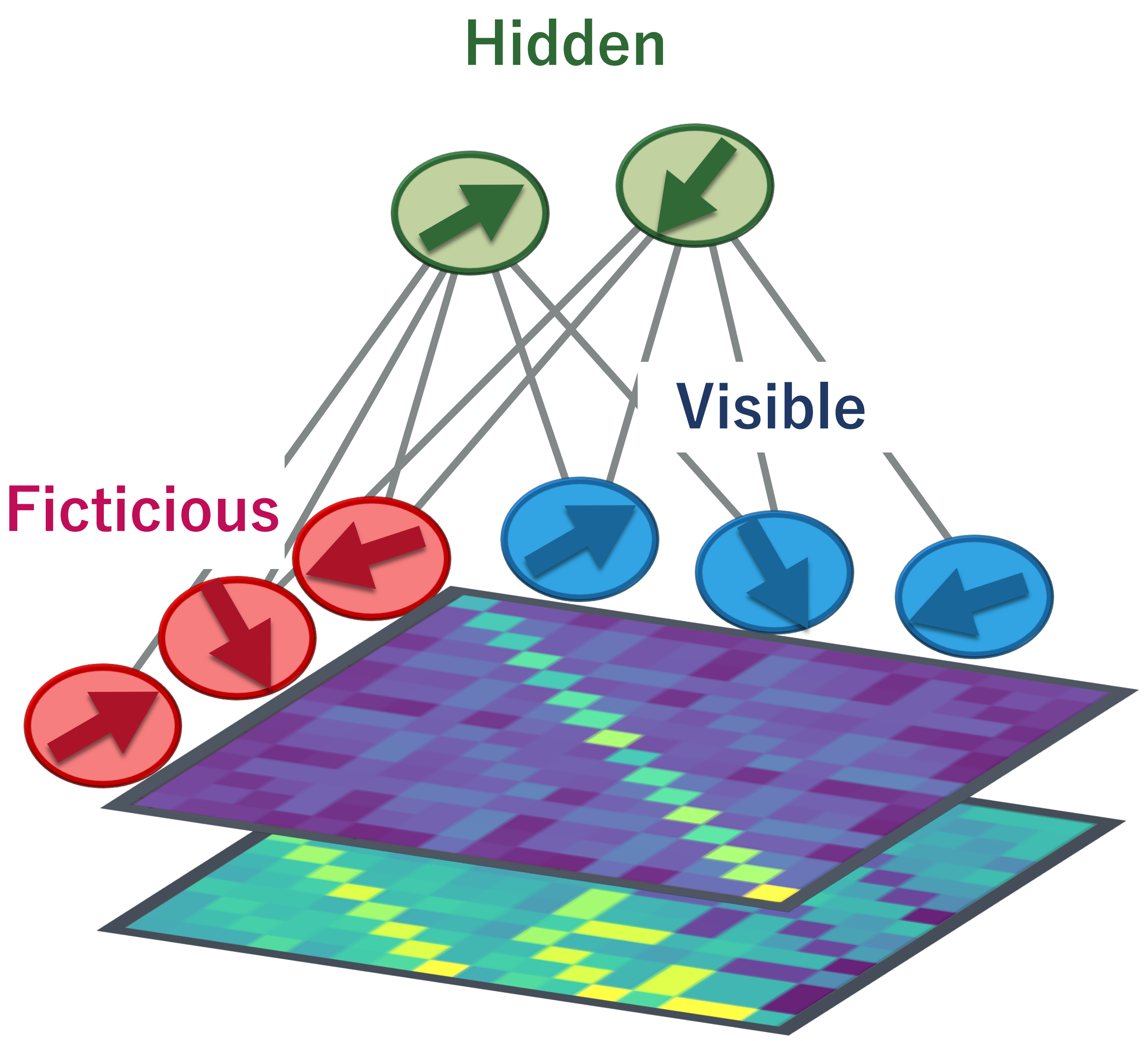
ニューラルネットワークで探る量子多体系の表現



吉岡信行（理研）

Zoom Webinar, 2020.07.09

自己紹介

吉岡　信行 (Nobuyuki Yoshioka)

2015.03　東京大学理学部物理学科　卒

2017.03　東大理物　修士課程修了 (桂研)

2020.03　東大理物　博士課程修了 (桂研)

博論タイトル「ニューラルネットワークによる物理状態の判定と表現」

現在　理化学研究所 開拓研究本部 Nori理論量子物理研究室

主な研究内容：量子物理・情報科学の境界から生まれる高速アルゴリズムの探索

・ニューラルネットワークによる量子多体状態の効率的表現

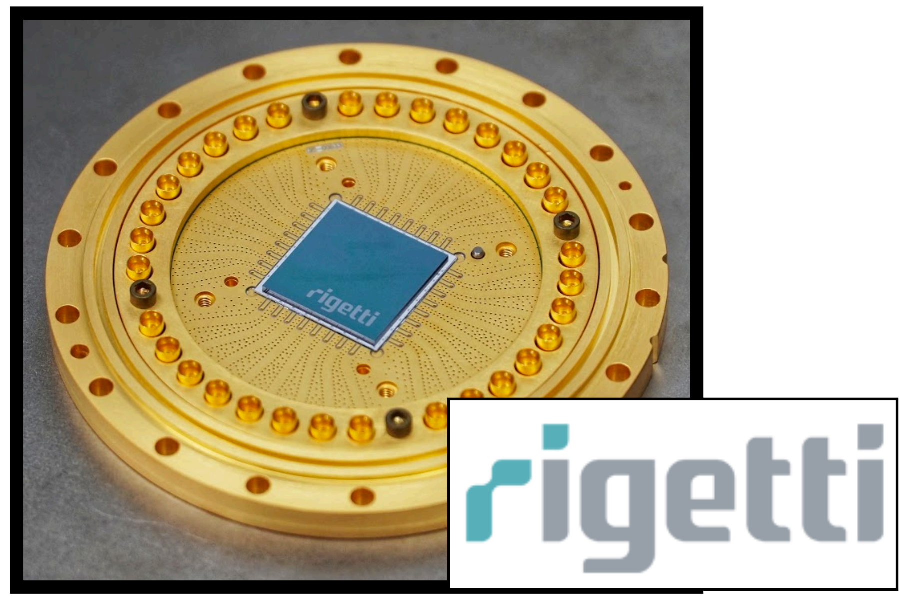
Today

・NIQSアルゴリズム開発

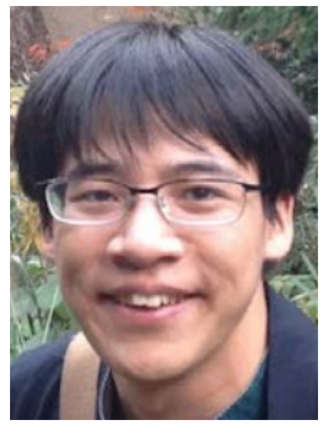
・非平衡量子ダイナミクス

e.g. 物理量のError bound

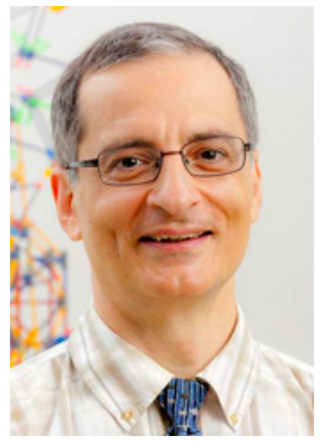
e.g.オンサーガー代数によるPerfect Scarの構築



共同研究者のみなさま



Dr. Ryusuke Hamazaki



Dr. Franco Nori

Dr. Ravindra Chhajlany

Dr. Clemens Gneiting

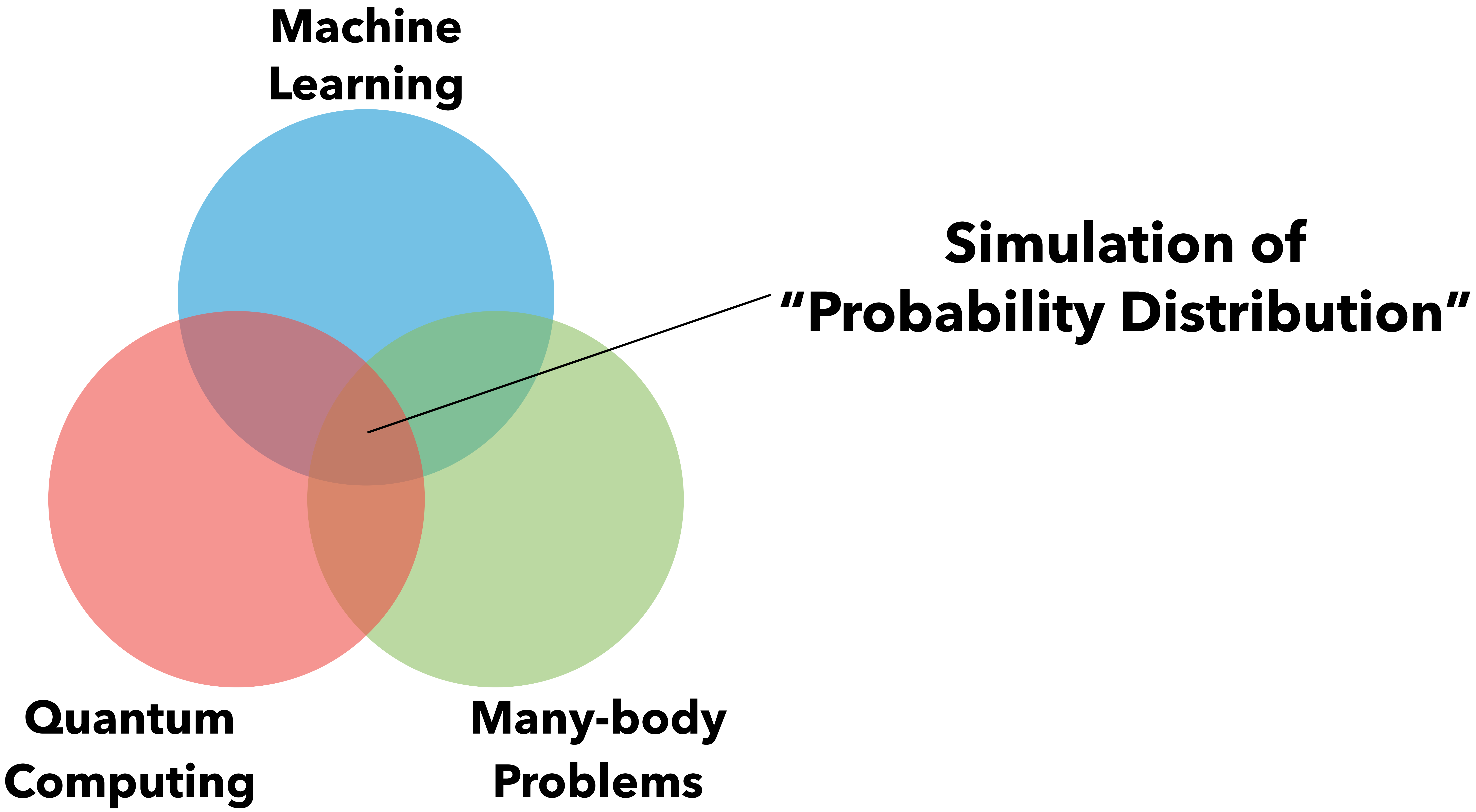
Prof. Wataru Mizukami

Introduction to Neural Quantum States

- Restricted Boltzmann Machine as a quantum state
- Variational calculation
- Relationship with tensor networks

Application to open quantum system

- Steady state as “ground state” of Lindbladian
- Results by RBM ansatz


Machine
Learning
Quantum
Computing
Many-body
Problems
Simulation of
"Probability Distribution"

Simulation of "Probability distribution"

	Data-driven learning Obtained via fitting into measurements No access to "ground truth"	**Model-driven learning** Obtained via optimization of model-based well-defined cost function
Classical data Probability distribution Single computational basis	**- Machine learning** e.g. image recognition e.g. data science	- Mathematical programming e.g. combinatorial optimization - Recommendation for Monte Carlo
Quantum data Wave function Unitary/non-unitary mapping	**- Quantum tomography** - Reproduce quantum state, Hamiltonian, Liouvilllian etc. **- Error mitigation**	**- Variational simulation** - Costless simulation of eigenvectors, density mat. - Tensor Nets, Neural Nets, Quantum circuits, etc.

Data-driven learning

Optimize the model p_λ s.t. "distance" between target P is minimized.

e.g. Kullback-Leibler divergence

$$D_{\mathrm{KL}}(P||p_\lambda) = \sum_{\mathbf{x}} P(\mathbf{x}) \ln\left(\frac{P(\mathbf{x})}{p_\lambda(\mathbf{x})}\right)$$

Variational simulation

Optimize the ansatz Ψ_θ w.r.t target depending on what you want to do

e.g. Hamiltonian

$$E_{\mathrm{vmc}} = \underset{\theta}{\mathrm{argmin}} \frac{\langle \Psi_\theta | H | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle}$$

e.g. Product of Liouvillian

$$0 = \underset{\theta}{\mathrm{argmin}} \frac{\langle\langle \rho_\theta | \hat{\mathcal{L}}^\dagger \hat{\mathcal{L}} | \rho_\theta \rangle\rangle}{\langle\langle \rho_\theta | \rho_\theta \rangle\rangle}$$

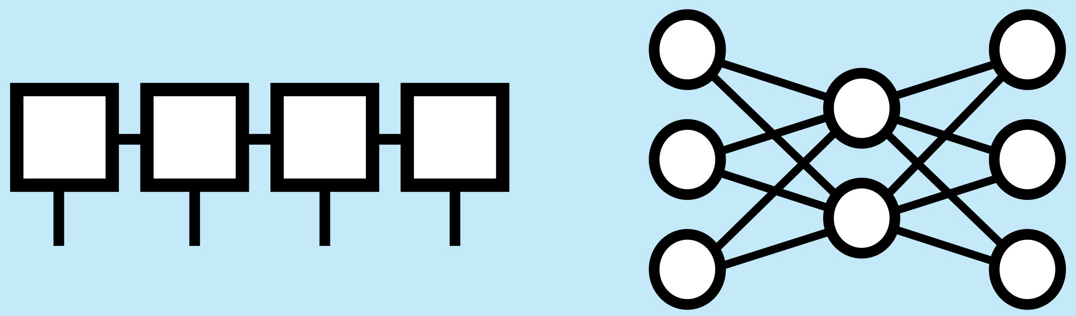
Model

$$H = \sum_q h_q P_q$$

"Ground truth" accessible by (super)-exponential cost

Minimize energy →

GS approximant Ψ_θ



Parametrized model

→ **Static property**
- Symmetry breaking
- Ordering

Minimize "energy" →

Steady state ρ_θ

→ **Non-equilibrium property**

Appropriate ansatz, "cost function", optimization needed

Simulation/Estimation of “Probability distribution”

	Data-driven Obtained via fitting into measurements No access to “ground truth”	**Model-driven** Obtained via optimization of model-based well-defined cost function
Classical data Probability distribution Single computational basis	**- Machine learning** e.g. image recognition e.g. data science	- Mathematical programming e.g. combinatorial optimization mendation for Monte Carlo
Quantum data Wave function Unitary/non-unitary mapping	**- Quantum tomography** - Reproduce quantum state, Hamiltonian, Liouvilllian etc. **- Error mitigation**	**- Variational simulation** - Costless simulation of eigenvectors, density mat. - Tensor Nets, Neural Nets, Quantum circuits, etc.

Neural Networks as variational ansatz

Carleo&Troyer ('17),

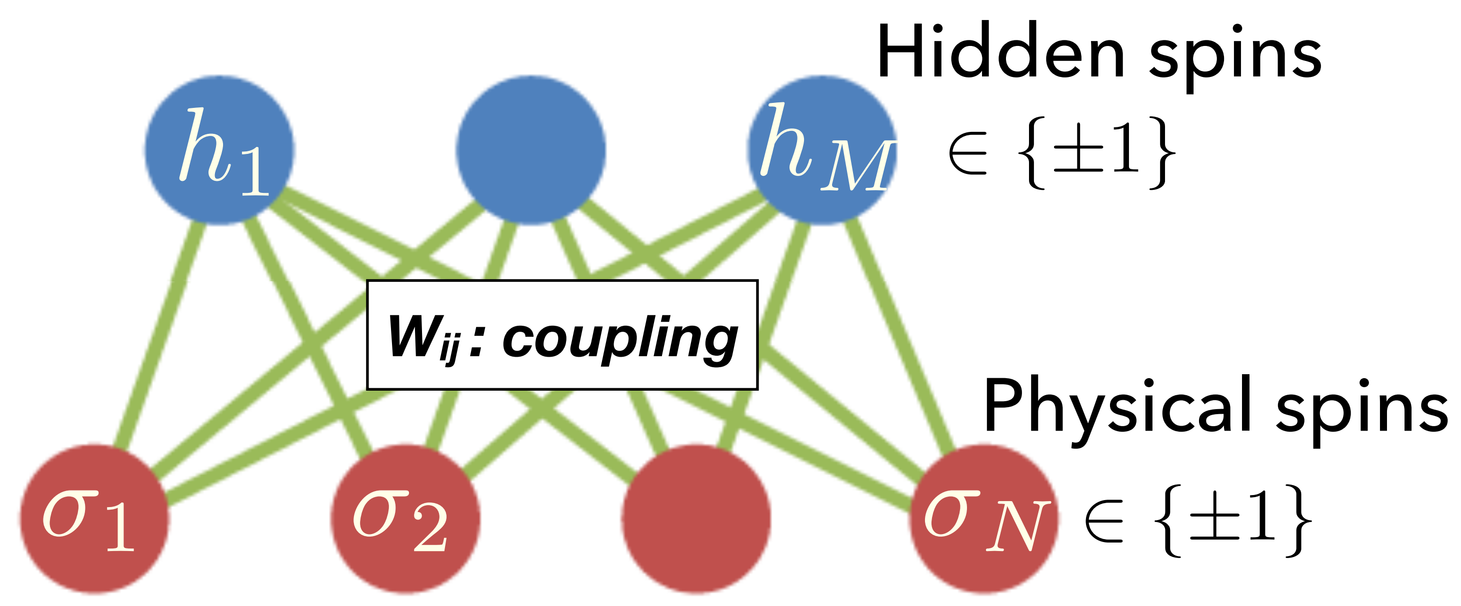
e.g. Restricted Boltzmann Machine

Dimension-free variational ansatz inspired by machine learning

$$|\Psi\rangle = \sum_{\sigma} \Psi(\sigma)|\sigma\rangle \quad \Psi(\sigma) \propto \sum_{h} e^{W_{ij}\sigma_i h_j + a_i\sigma_i + b_j h_j}$$

Interaction ($W_{ij}\sigma_i h_j$), Mag. fields ($a_i\sigma_i + b_j h_j$)

h ← Tracing out aux. space



Restricted Boltzmann Machine

Optimization by variational Monte Carlo

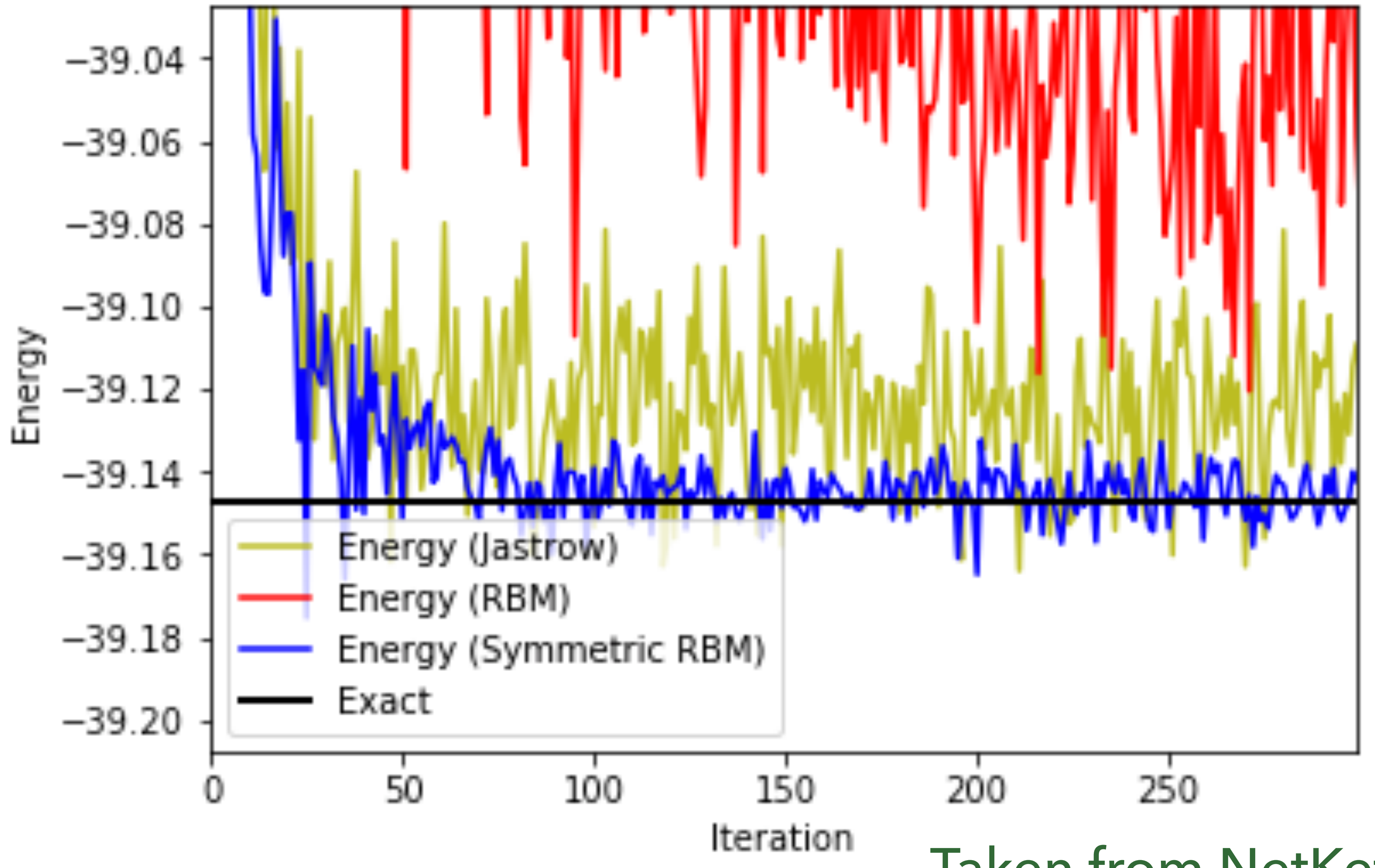
Optimize a,b,W that minimize the GS energy,

$$E_{GS} = \underset{\Psi}{\text{argmin}} \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}$$

e.g. stochastic gradient descent,

$$a_i \to a_i - \eta\partial_{a_i}\langle H\rangle$$

GS energy optimization of 1d AF Heisenberg (L=22)



Taken from NetKet website

RBM as ground state ansatz

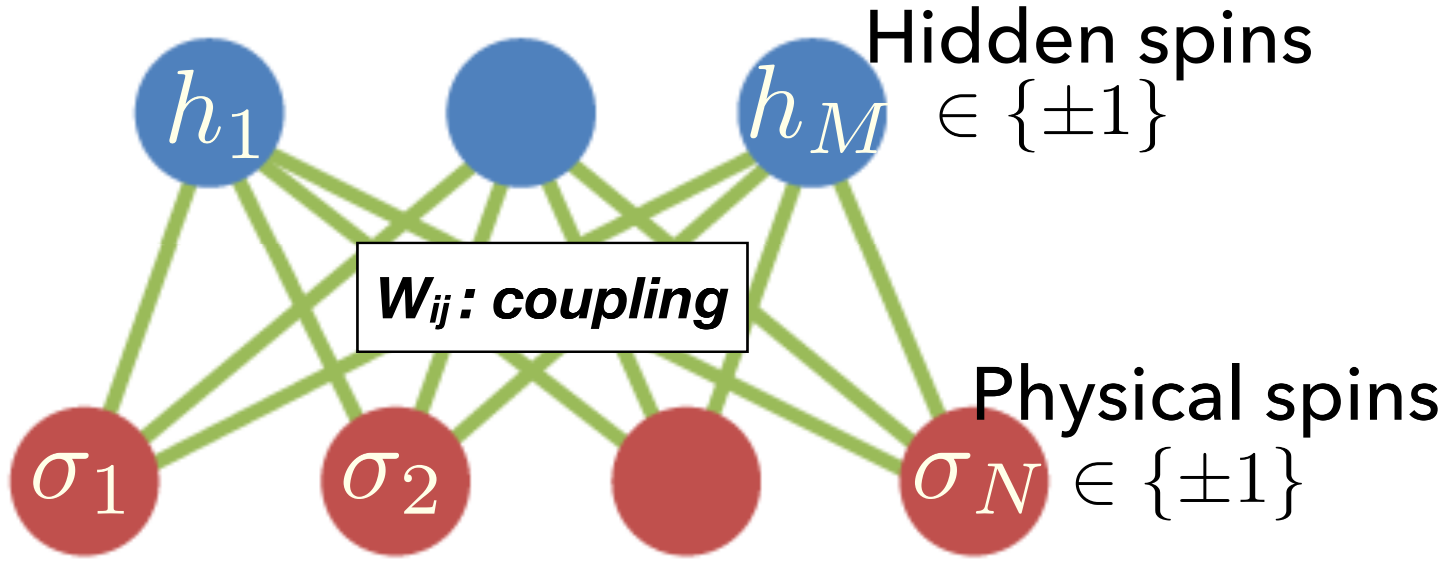
e.g. Restricted Boltzmann Machine

Dimension-free variational ansatz inspired by machine learning

Interaction: $W_{ij}\sigma_i h_j$ Mag. fields: $a_i\sigma_i + b_j h_j$

$$|\Psi\rangle = \sum_{\sigma} \Psi(\sigma)|\sigma\rangle \quad \Psi(\sigma) \propto \sum_{h} e^{W_{ij}\sigma_i h_j + a_i\sigma_i + b_j h_j}$$

h ← Tracing out aux. space



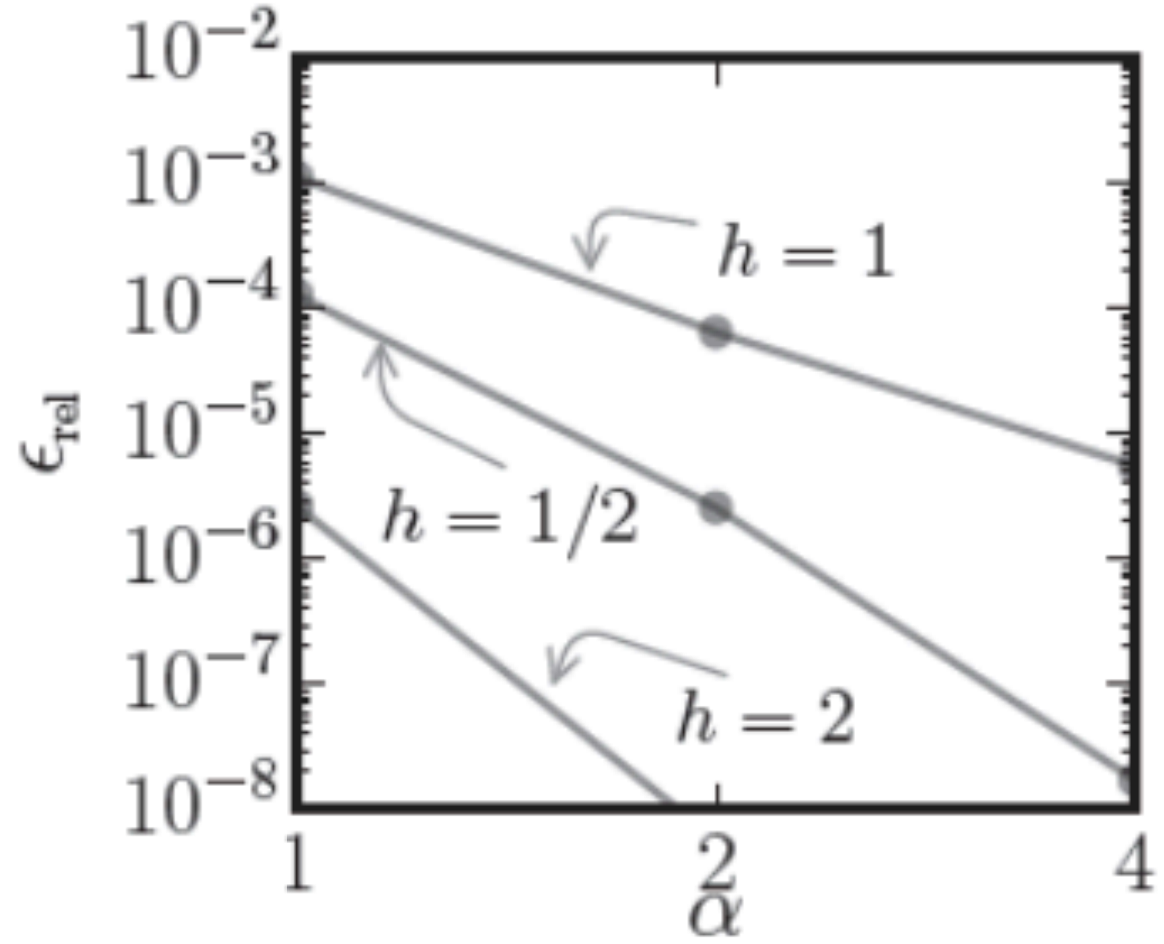
Restricted Boltzmann Machine

Comparison with other ansatz

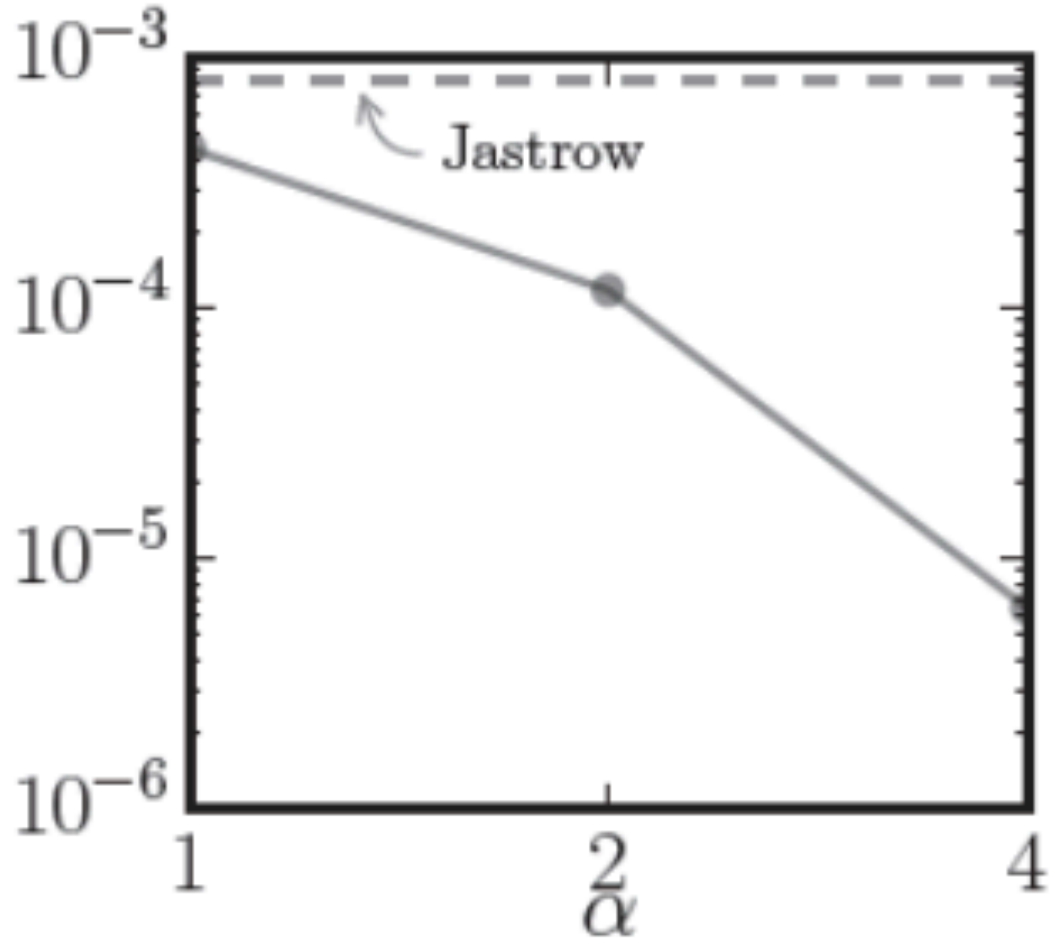
Carleo&Troyer Science 355('17)

State-of-art GS energy accuracy achieved by variational imaginary-time evolution

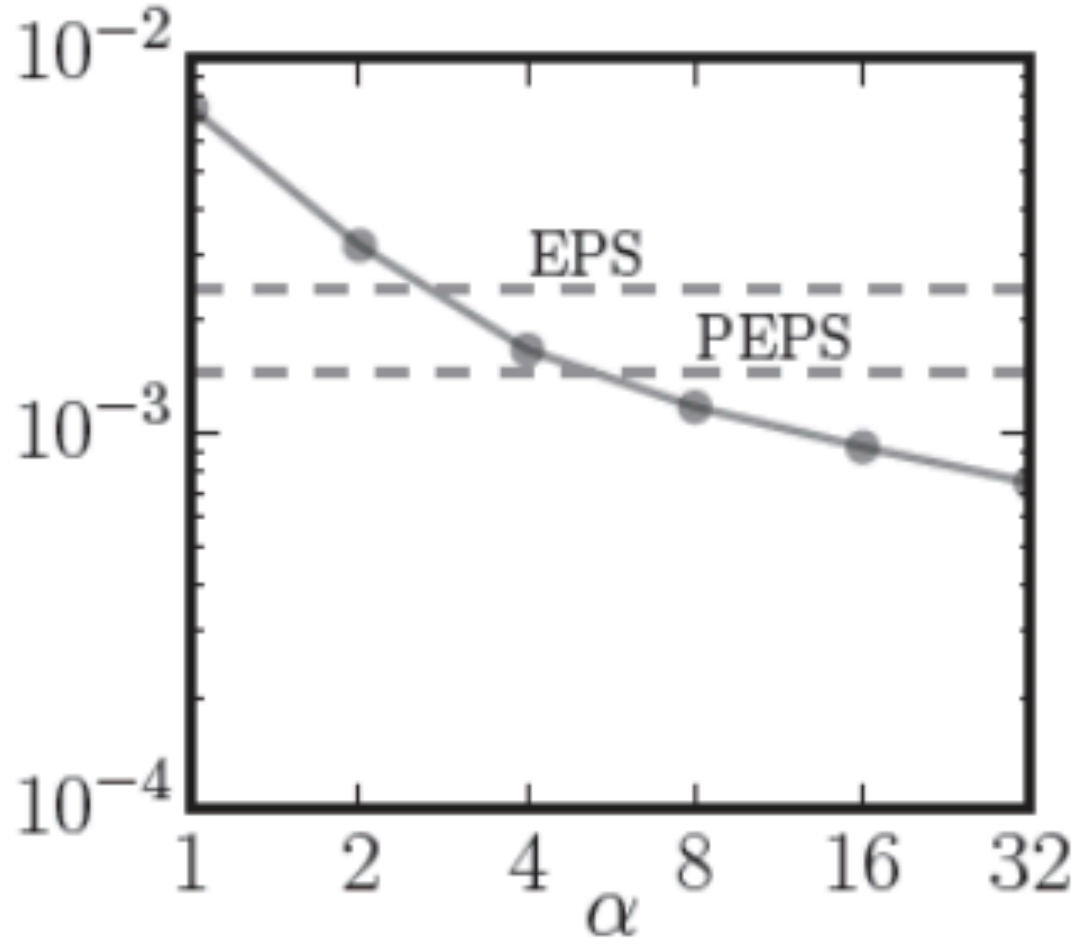
1d Traverse-field Ising model
80 spins, periodic boundary,
h : field, alpha: (# of hidden neuron)



1d AF Heisenberg model
80 spins, periodic boundary



2d AF Heisenberg model
10x10 spins, periodic boundary



About RBM-type ansatz

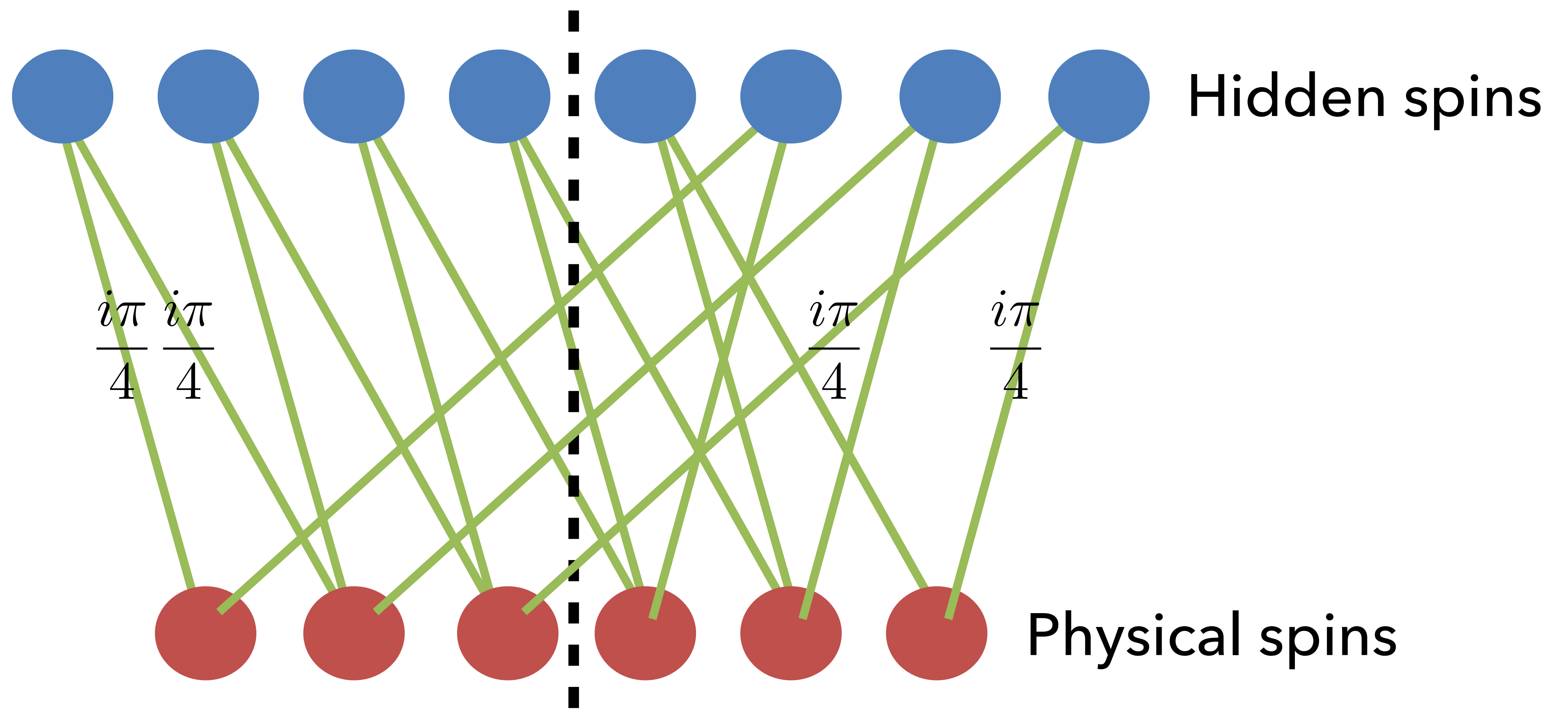
Entanglement property

Deng etal. PRX ('17)

e.g.) Maximized half-chain Renyi entropy for 2n spins with $\mathcal{O}(n)$ parameters

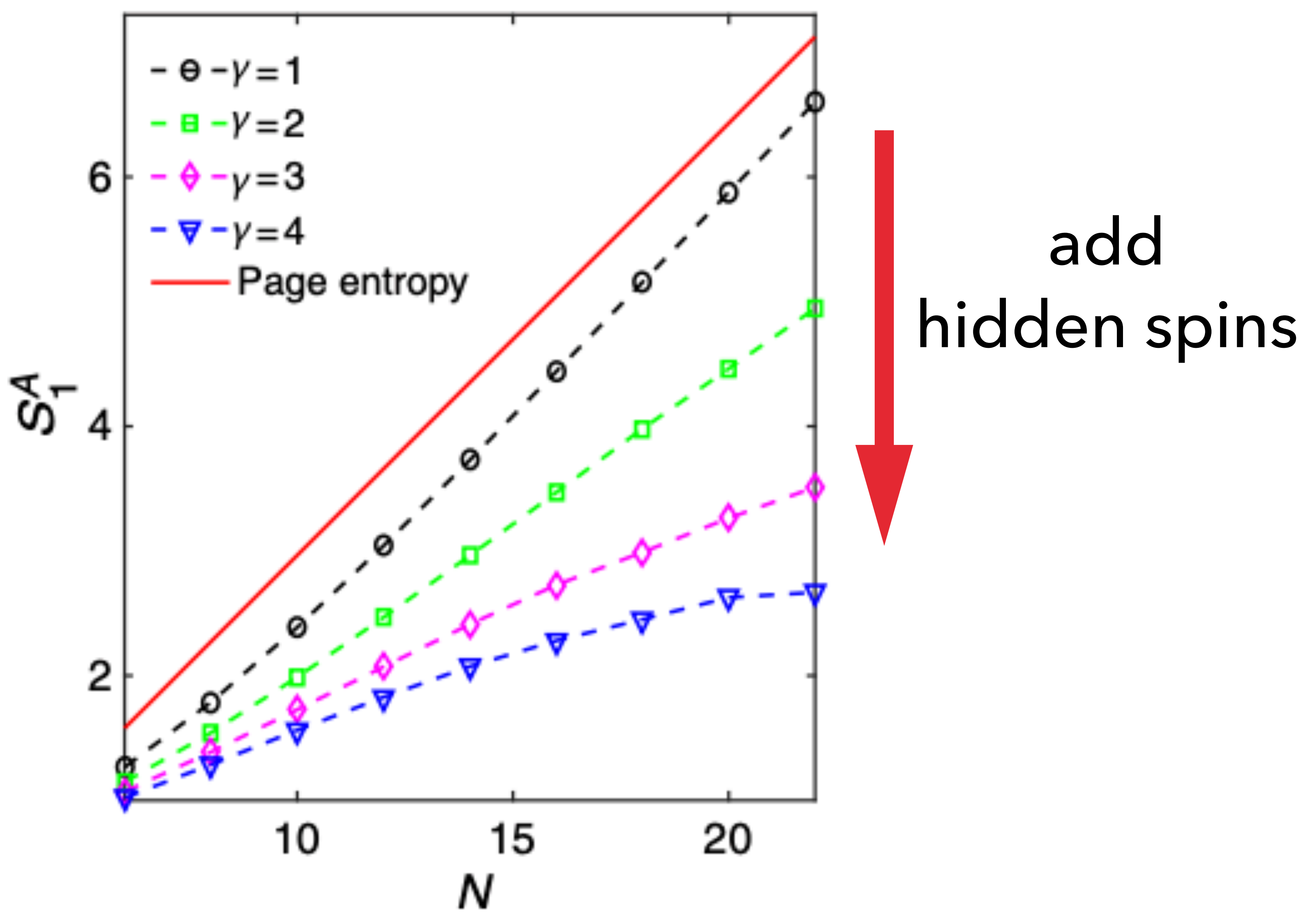
$$S_\alpha^A = n \log 2, \forall \alpha$$

$$S_\alpha^A \equiv \frac{1}{1-\alpha} \log[\mathrm{Tr}(\rho_A^\alpha)]$$



e.g.) Random RBM states

von Neumann Entanglement scaling



- EE(RBM) < Page value (average EE of Haar random)
  → Not all random states efficiently captured

- In fact you need exponentially many parameters to fit Haar random states

- Level spacing of entanglement ham obeys Poisson (expected to correspond to those of integrable system)

Other types of neural quantum state

Feed-forward neural nets

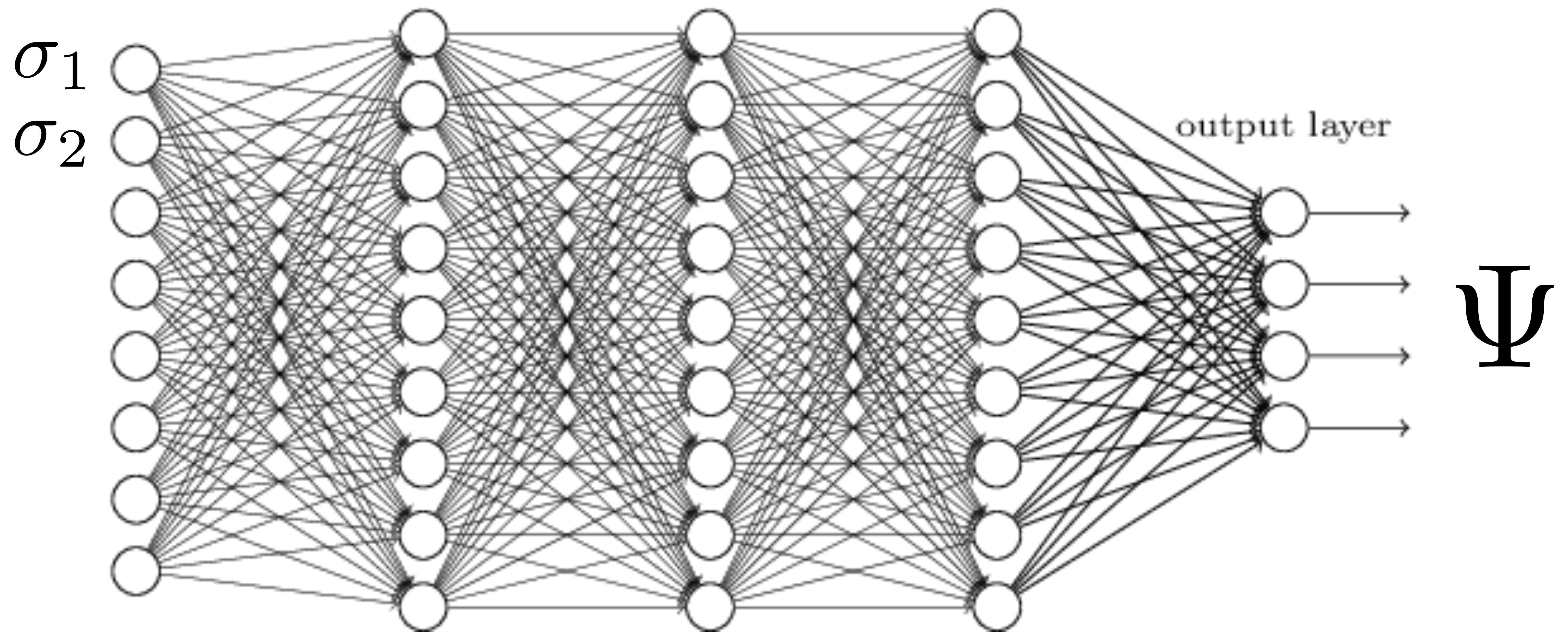
non-linear

$$\Psi(\sigma) \propto \mathcal{N}_n\mathcal{L}_n \cdots \mathcal{N}_2\mathcal{L}_2\mathcal{N}_1\mathcal{L}_1(\sigma)$$

linear

Intermediate output given as

$$\mathbf{y}_n = \mathcal{N}_n\left(\mathcal{L}_n(\mathbf{y}_{n-1})\right)$$



e.g. Fully-connected Saito ('17), Cai&Liu ('18), Choo et al. ('17)

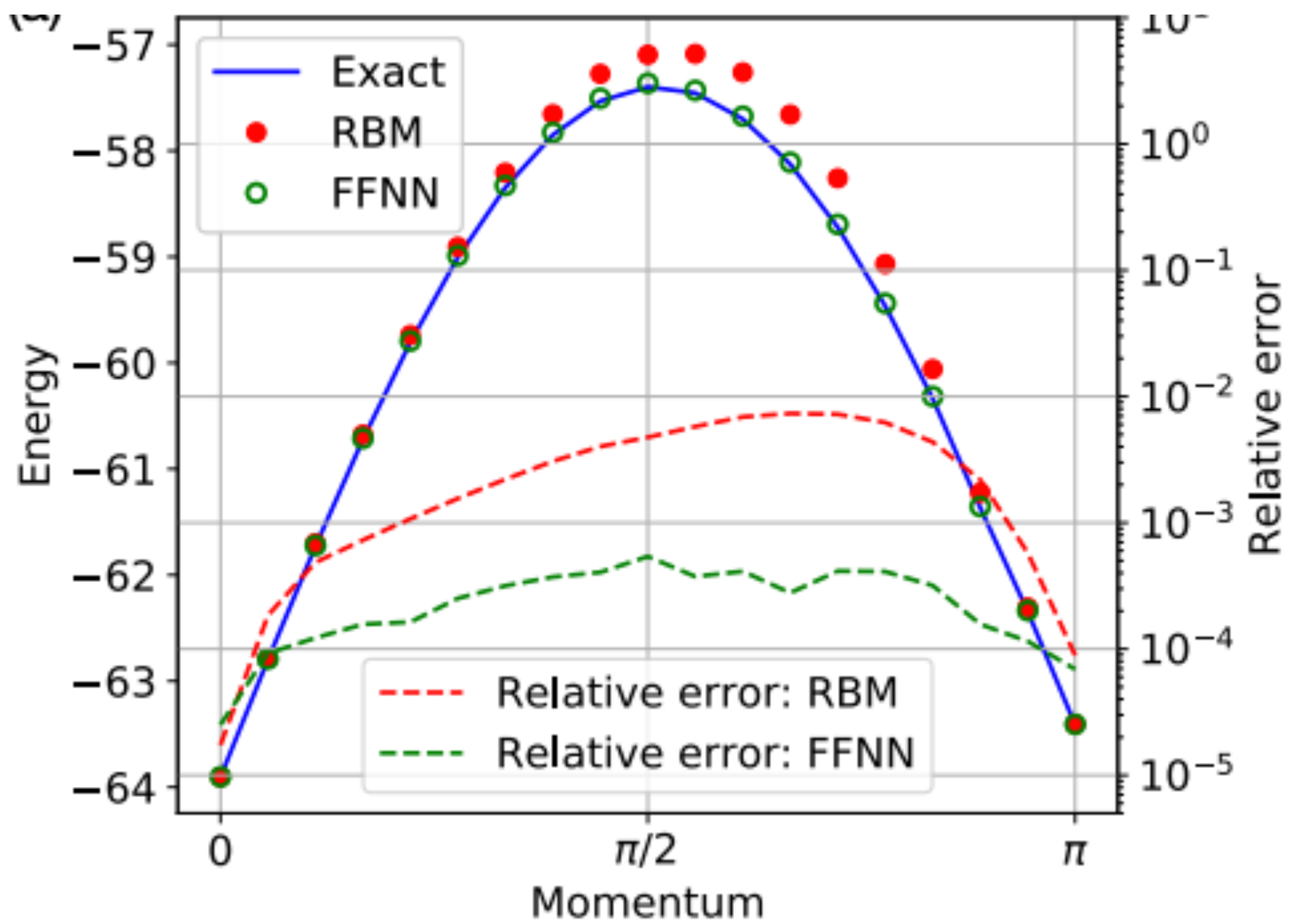
$$\begin{cases} \text{Linear: } (\mathcal{L}(\sigma))_j = \sum_i W_{ij}\sigma_i + b_j \\ \text{Non-linear: } \mathcal{N}(z_j) = \tanh(z_j) \end{cases}$$

- Better performance in excited states for 1d AFH

- Entanglement spreading is slow, not so powerful in d>1 systems

Momentum-resolved energy

1dAFH model (L=36 sites) Choo et al., PRL ('17)

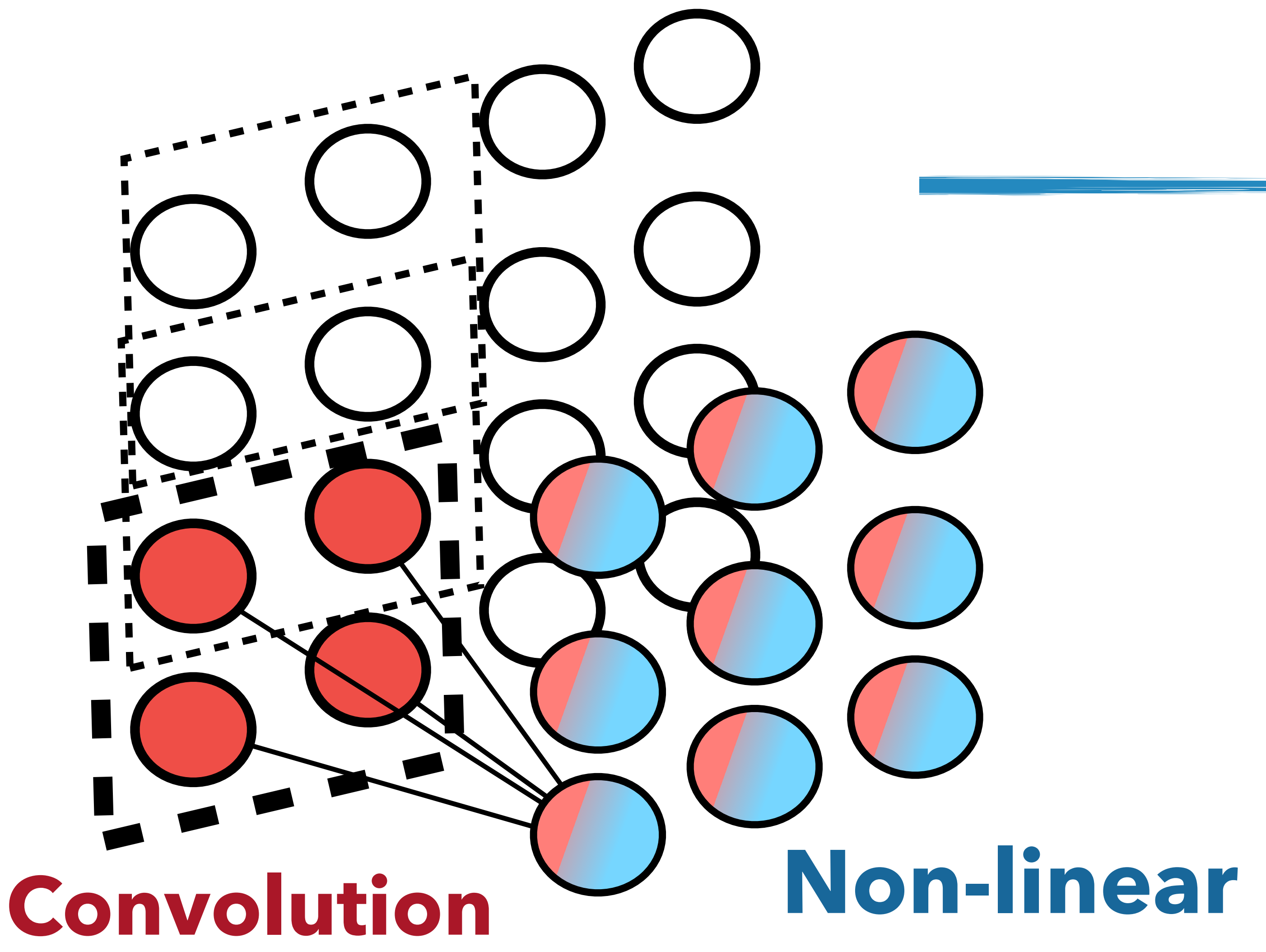


Other types of neural quantum state

Feed-forward neural nets

$$\Psi(\sigma) \propto \mathcal{N}_n \mathcal{L}_n \cdots \mathcal{N}_2 \mathcal{L}_2 \mathcal{N}_1 \mathcal{L}_1(\sigma)$$

non-linear: $\mathcal{N}_1$; linear: $\mathcal{L}_1$



e.g. Convolutional neural nets Choo et al., ('19), Szabo&Castelnovo ('20)

- Local receptive fields acting on finite regions

- MERA-like structure with higher efficiency for volume-law entanglement Levine et al., PRL ('19)

  RBM: $O(N)$ parameters in 2d

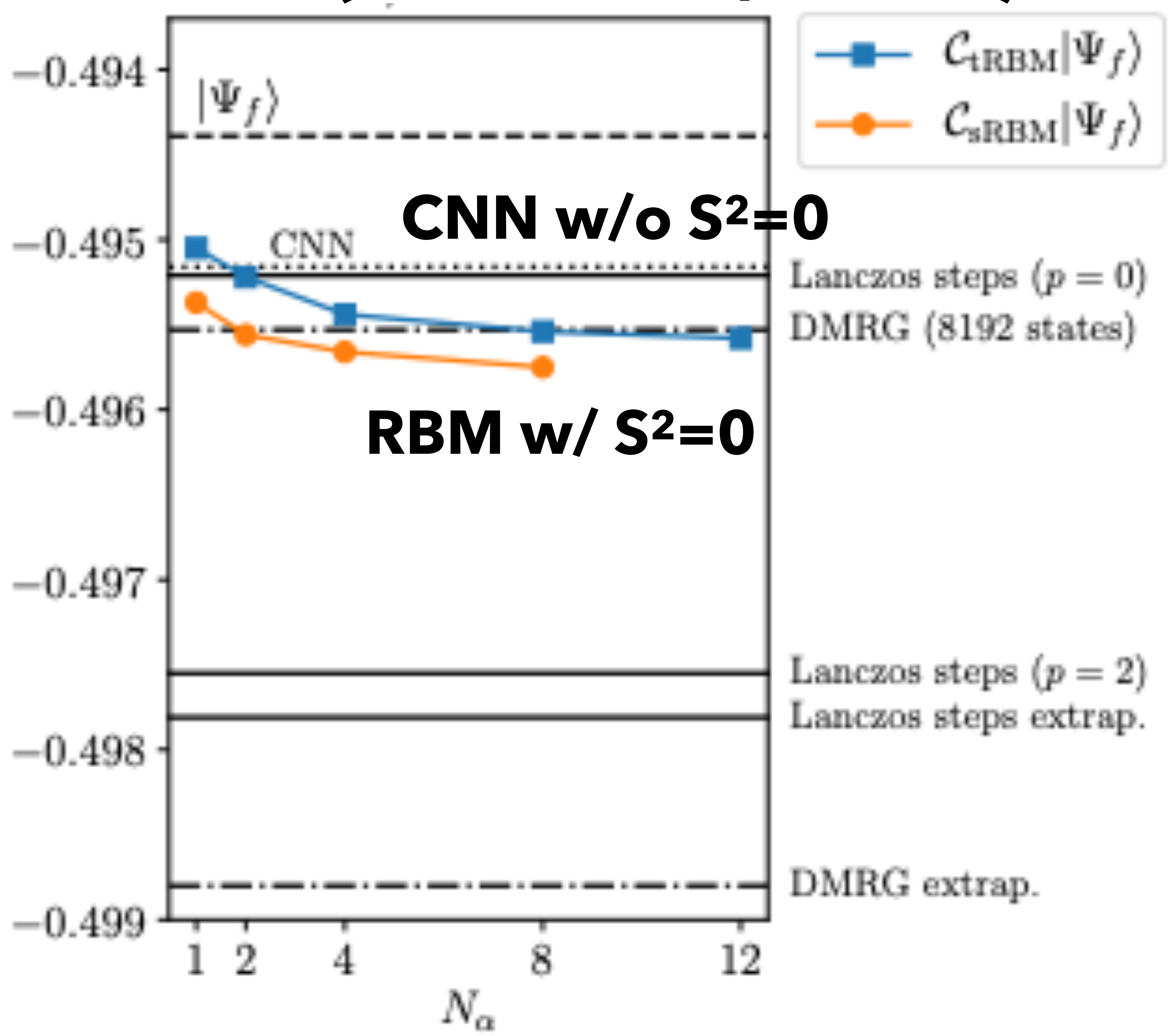
  CNN: $O(\sqrt{N})$ parameters in 2d

- Performance not the best without implementing symmetry (e.g. 2D J1-J2 Heisenberg model)

  Ferrari et al., ('19) Nomura&Imada ('20)

Ground-state energy

J1J2 model (J2/J1 = 0.5, 10x10) Ferrari et al., ('19)

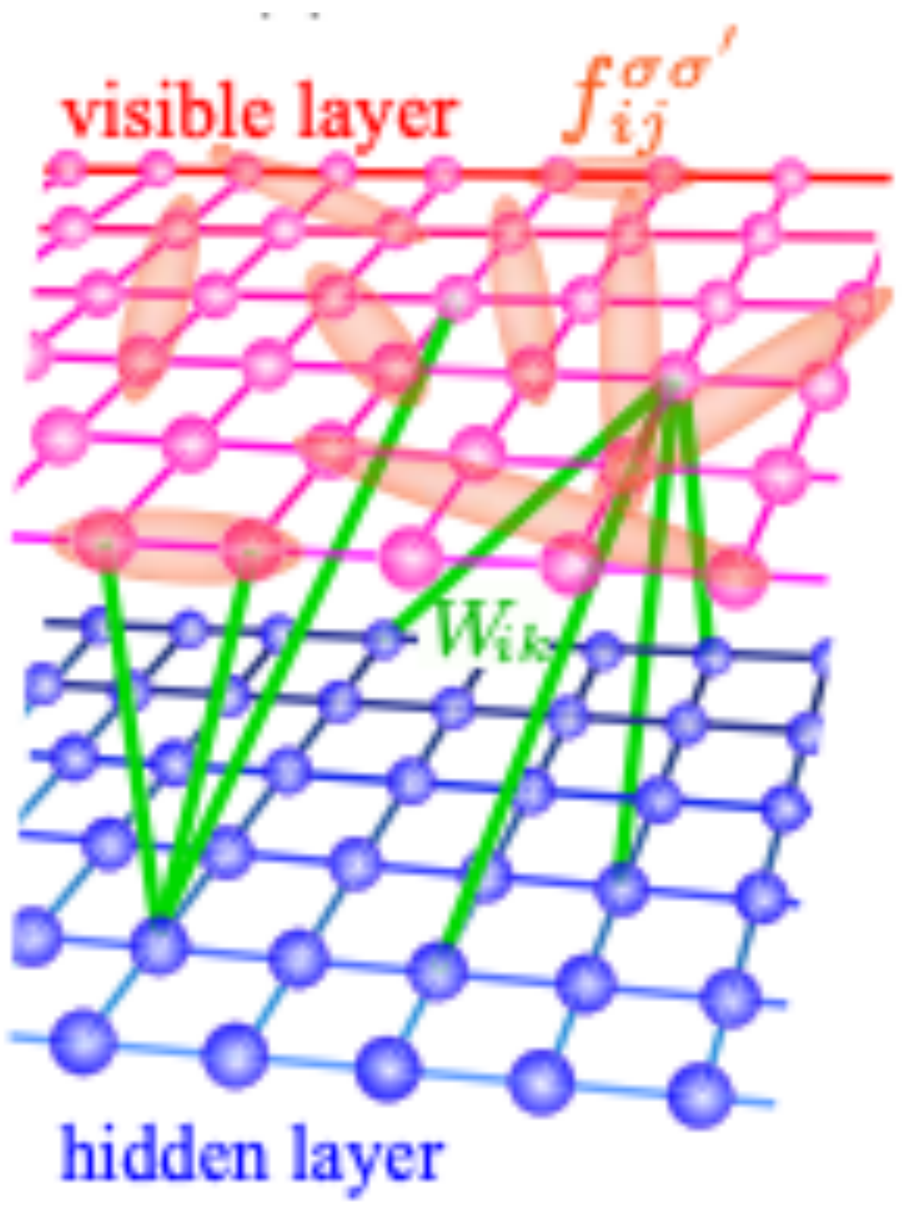


Other types of neural quantum state

Combination with conventional technique Nomura et al., (2017)

$$\Psi(\sigma) = \Psi_{\mathrm{RBM}}(\sigma)\psi_{\mathrm{PP}}(\sigma)$$

Pair product
originally introduced to study fermonic system

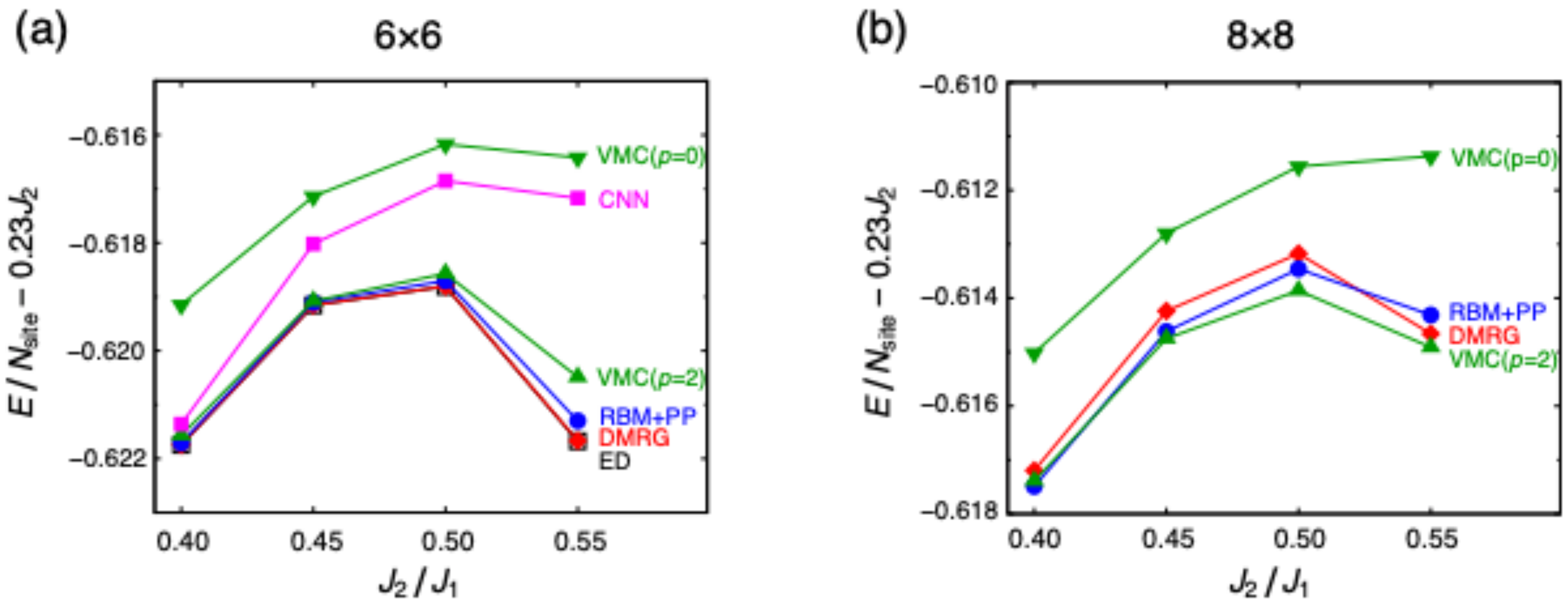


Based on two-body interaction between visible spins:

$$|\psi_{\mathrm{PP}}\rangle = \sum_{\sigma} \psi_{\mathrm{PP}}(\sigma) \left(\prod_i c_{i\sigma}^{\dagger} |0\rangle \right) = P_G \left(\sum_{i,j} f_{i,j}^{\uparrow\downarrow} c_{i\uparrow}^{\dagger} c_{j\downarrow}^{\dagger} \right)^{N_{\mathrm{site}}/2} |0\rangle$$

P_G: prohibits double occupancy

Result for J1J2 Heisenberg model Nomura & Imada (2020)



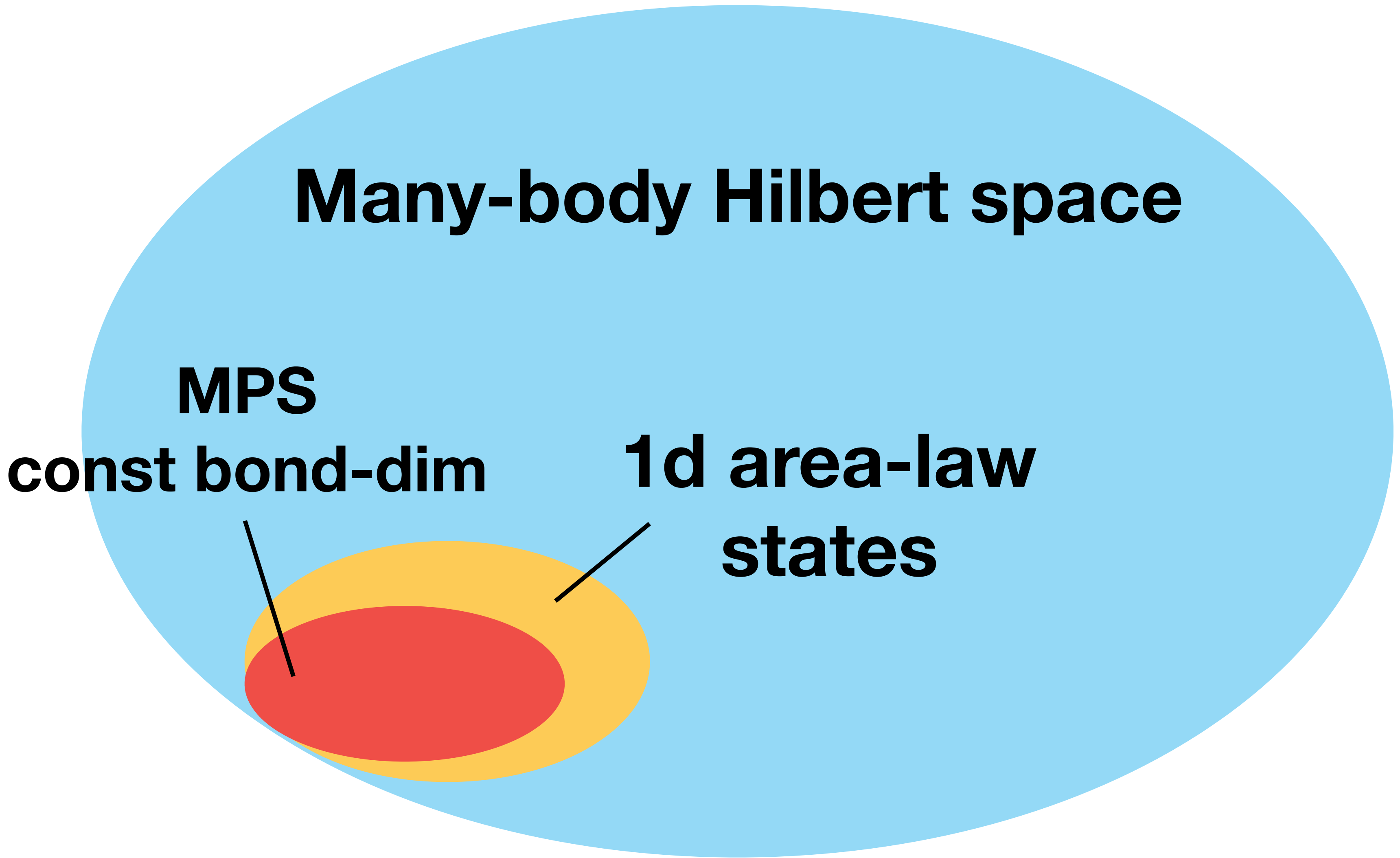
Relation with other variational ansatz

Tensor networks

- Class of variational ansatz based on tensor contraction

$$|\Psi\rangle = \sum_{\sigma} \mathrm{Cont}\left(\bigotimes_{\alpha} T_{\alpha}\right)_{\sigma} |\sigma\rangle$$

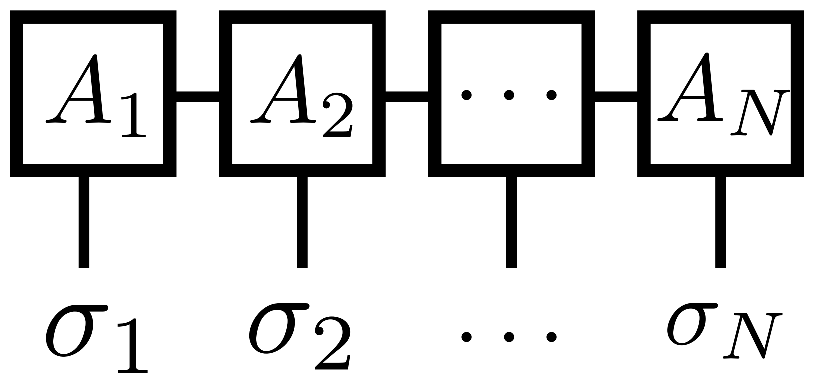
- Advantageous especially for area-law entangled states



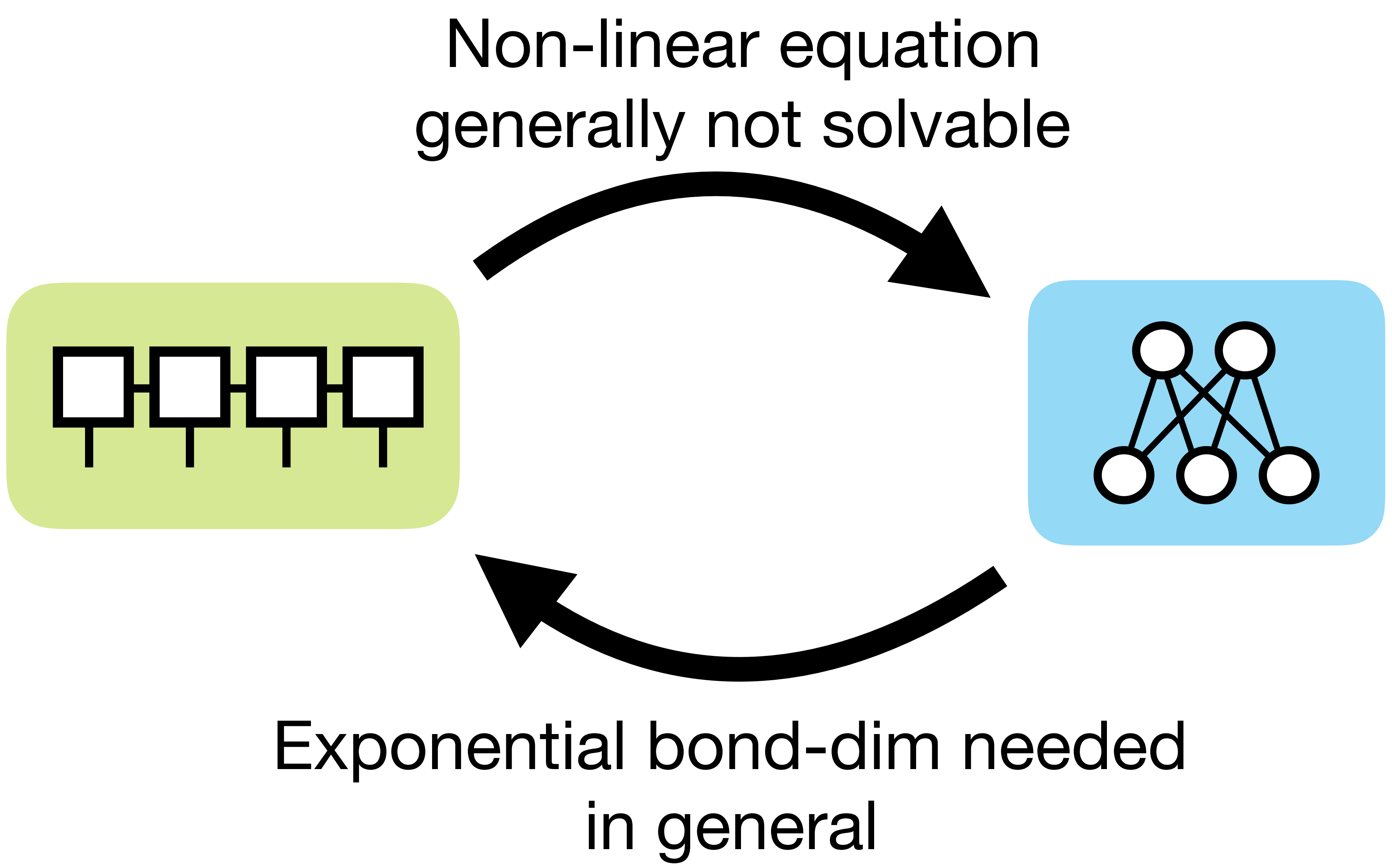
e.g. Matrix Product States Fannes et al. ('92), White ('92),

$$|\Psi\rangle = \mathrm{Tr}\left(\prod_{i} A_i^{[\sigma_i]}\right) |\sigma\rangle$$

- Killer app in gapped 1d when optimized via DMRG
- Not scalable for generic d>2 systems



MPS ⇆ RBM is hard Chen et al. ('18),



Relation with other variational ansatz

RBM as subclass of "tensor network": String Bond States Glasser etal. PRX ('18)

- Product of MPSs in subset of system

$$|\Psi\rangle = \prod_S \mathrm{Tr}\left(\prod_{i\in S} A_{i,S}^{[\sigma_i]}\right)|\sigma\rangle$$

(cf. MPS
$$|\Psi\rangle = \mathrm{Tr}\left(\prod_i A_i^{[\sigma_i]}\right)|\sigma\rangle$$
)

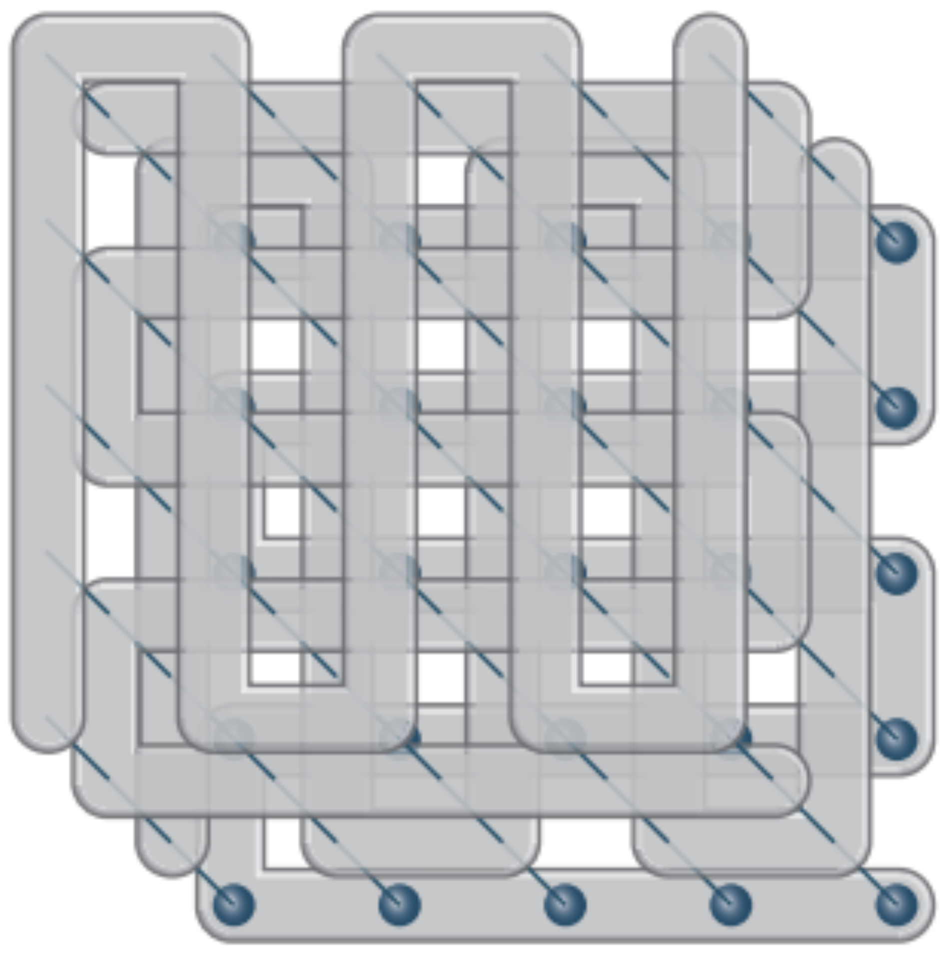
RBM, defined as follows, can also be written in the above form

$$\Psi(\sigma) \propto \sum_{h_j=\pm 1} \exp\left(\sum_{ij} W_{ij}\sigma_i h_j + \sum_j b_j h_j\right)$$

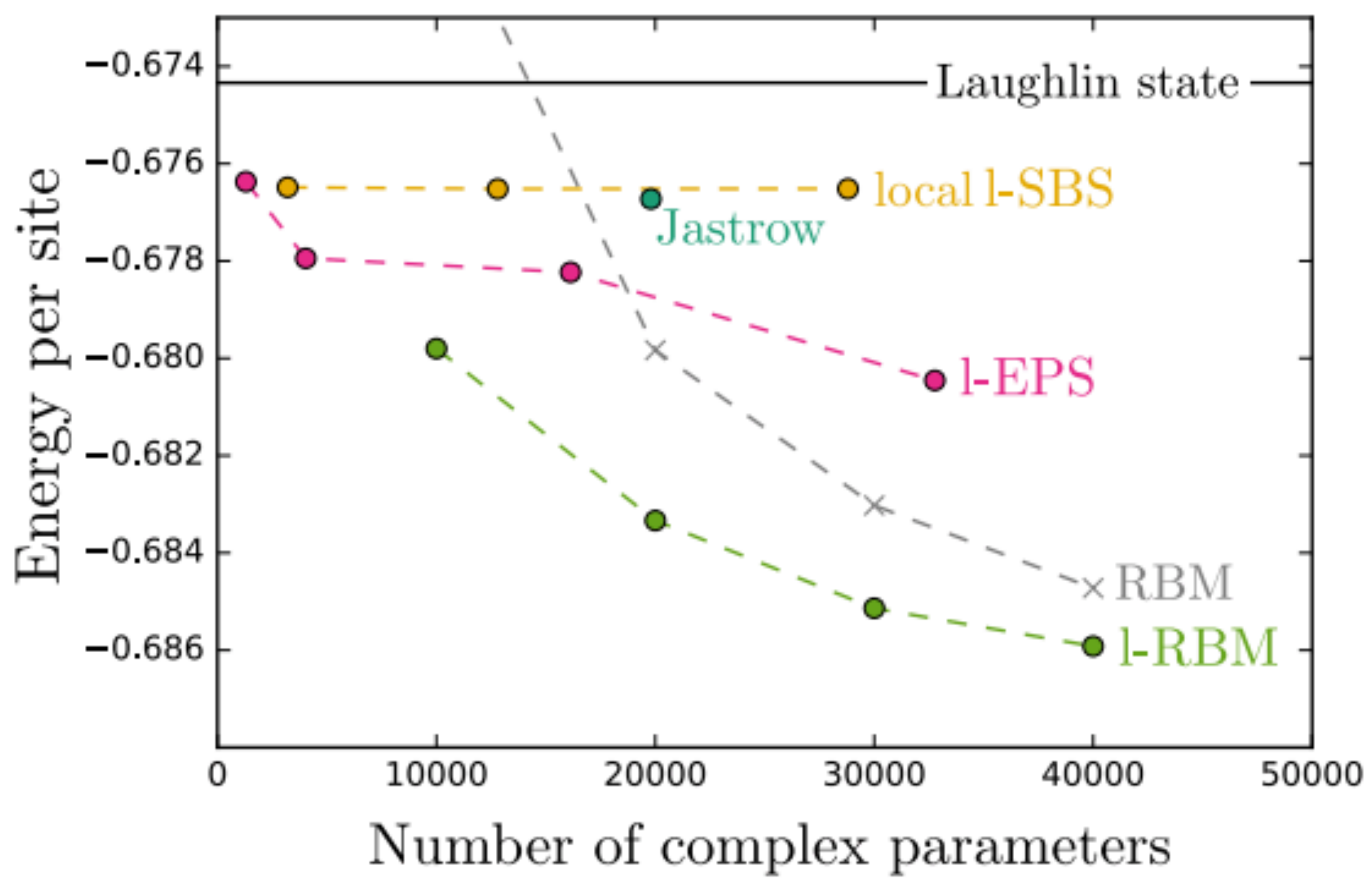
- Very generic, highly-dependent on "string" choices.

Equivalent to single MPS with exponential bond dimension

- DMRG not available but allows Monte Carlo sampling Schuch etal. PRL ('08)



(b) RBM as a non-local SBS



(c) 10x10 lattice

Variational simulation by neural networks

Algorithm	Ansatz
Ground state (condensed matter) Carleo&Troyer Science ('17) Cai&Liu, PRB ('18) Nomura etal. PRB ('17) Glasser etal. PRX ('18)	**Convolutional neural nets** Choo etal. PRB ('19) Szabo&Castelnovo ('20)
Ground state (quantum chemistry) Choo et al. ('20) Pfau et al. ('19) Hermann et al. ('19) **Yoshioka, Mizukami, Nori, in prep.** **Tomorrow 15:30 @FQCS2020**	**Autoregressive models** Levine etal. PRL ('20) Davis&Wang ('20) Hibat-Allah et al. ('20)
Real Time evolution Carleo&Troyer ('17) Czischek etal. PRB ('17) Schmitt&Heyl ('19) Lopez-Gutierrez&Mendl ('19)	**Volume-law entanglement in NQS** Deng et al. PRX ('17) Levine etal. PRL ('19)
Steady State in open system **Yoshioka&Hamazaki PRB ('19)** Nagy&Savona PRL ('19) Hartmann&Carleo PRL ('19) Vincenti etal. PRL ('19)	

Disclaimer: Table not exclusive

Introduction to Neural Quantum States

- Restricted Boltzmann Machine as a quantum state
- Variational calculation
- Relationship with tensor networks

Application to open quantum system

- Steady state as “ground state” of Lindbladian
- Results by RBM ansatz

Open quantum system

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

System **w/ dissipation**

Microscopic & coherent

e.g.) Superconducting qubits

e.g.) Cold atom, Trapped ions

interact

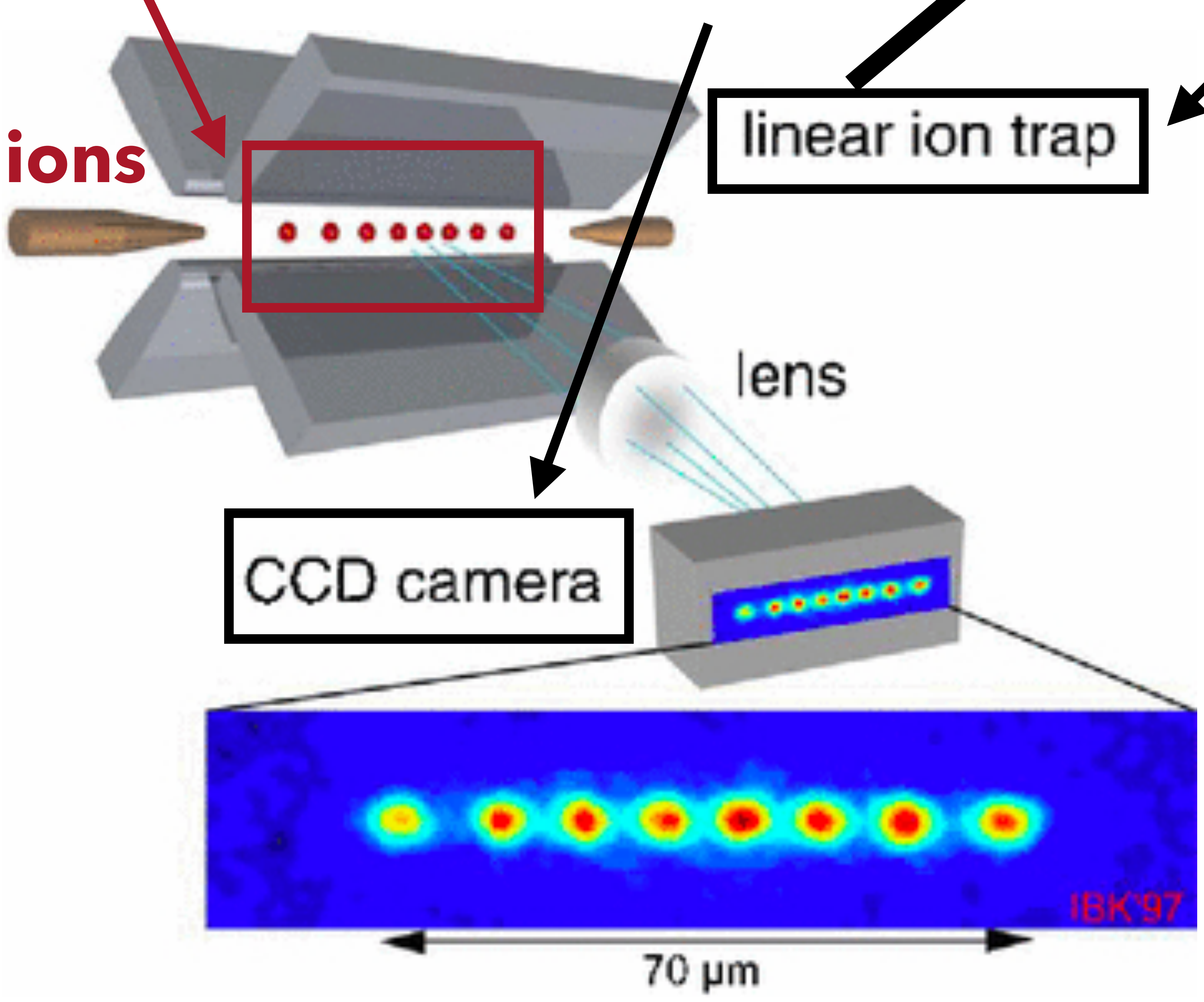
Environment

Ma ke

e. nent

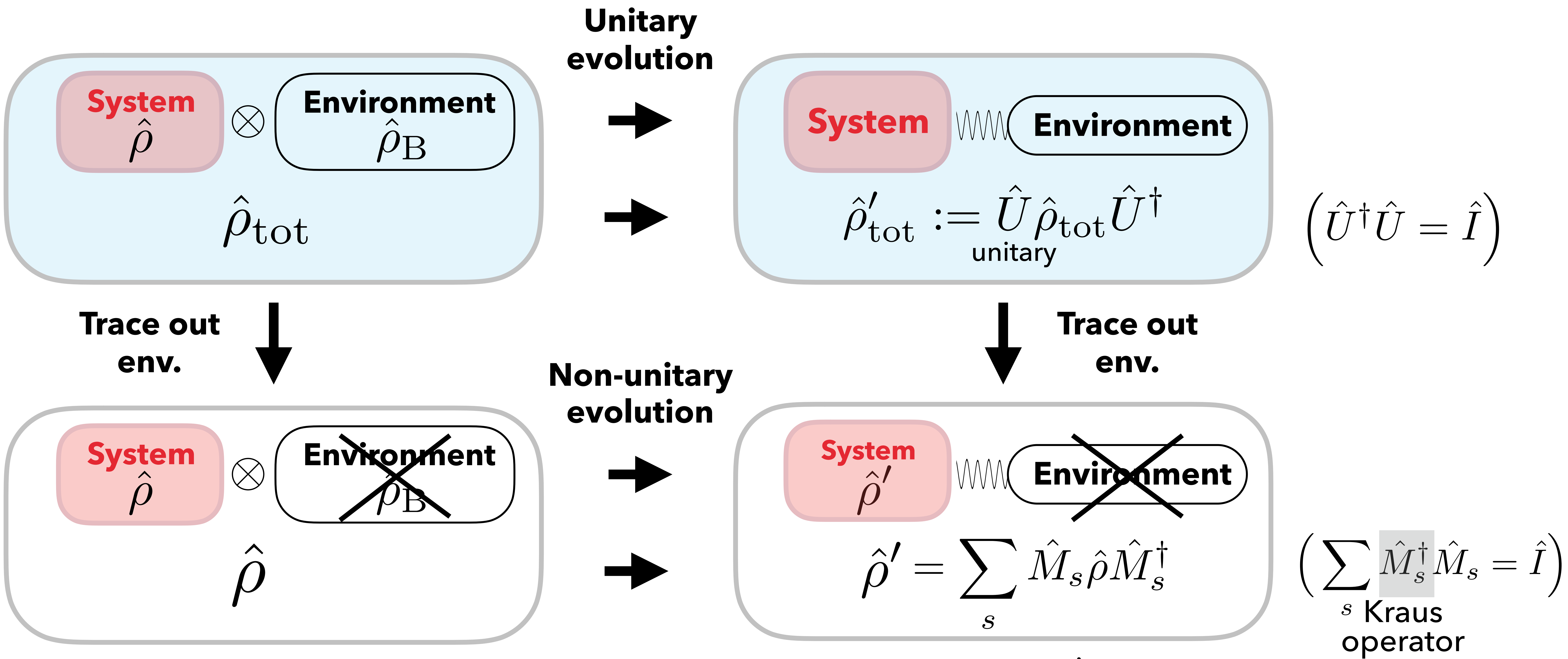
Trace out

e.g.) Laser shining



Taken from Garcia-Repoll, J.Phys B ('05)

Setup: Time evolution in open quantum system



$$\mathrm{Tr}[\hat{\rho}] = \mathrm{Tr}[\hat{\rho}'] = 1$$

Non-unitary but
Trace-preserving map

GKSL formalism and stationary state

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

Impose quantum master equation to satisfy

1. Completely positive **and** trace preserving (CPTP)
2. Time-locality (Markovianity, short correlation time in env.)

→ **Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) equation** Lindblad (1976) Gorini, Kossakowski&Sudarshan (1976)

$$\frac{d\hat{\rho}(t)}{dt} = \mathcal{L}(t)\hat{\rho}(t) := -i[\hat{H}(t), \hat{\rho}(t)] + \mathcal{D}(t)[\hat{\rho}(t)]$$

Liouvillian; Unitary dynamics; Non-unitary, but trace-preserving

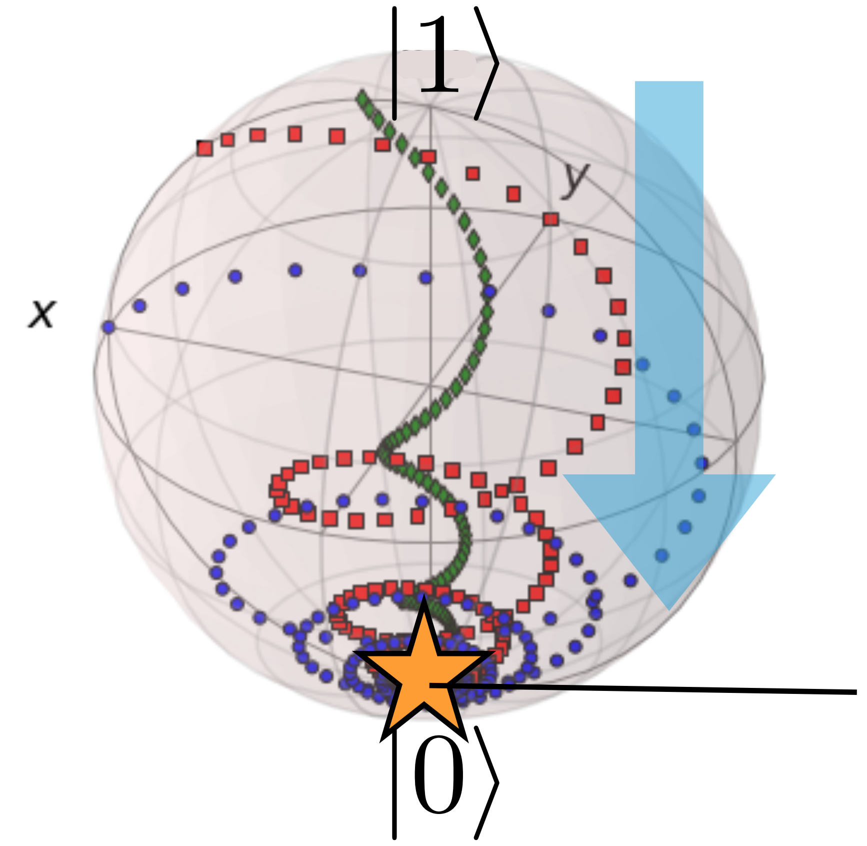
$$\mathcal{D}(t)[\hat{\rho}(t)] = \sum_s \hat{\Gamma}_s(t)\hat{\rho}(t)\hat{\Gamma}_s^\dagger(t) - \frac{1}{2}\left\{\hat{\Gamma}_s^\dagger(t)\hat{\Gamma}_s(t), \hat{\rho}(t)\right\}$$

If $\hat{H}(t) = \hat{H}$ and $\mathcal{D}(t) = \mathcal{D}$, i.e., time-homogeneous,

at least one non-equilibrium stationary state assured

Solution for $\frac{d\hat{\rho}_{\mathrm{SS}}}{dt} = 0$

e.g. Two-lev. system with damping



Variational Search for stationary states

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

Step1

Lindblad eqn. in vector representation

$$\frac{d|\rho(t)\rangle\rangle}{dt} = \hat{\mathcal{L}}|\rho(t)\rangle\rangle$$

$$|i\rangle\langle j| \mapsto |i\rangle \otimes |j\rangle$$

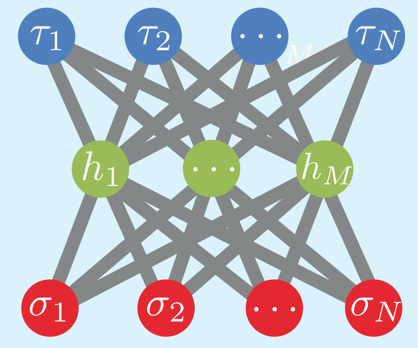
Step2

Define optimization task

$$\arg\min_{\rho} \frac{\langle\langle\rho|\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}|\rho\rangle\rangle}{\langle\langle\rho|\rho\rangle\rangle}$$

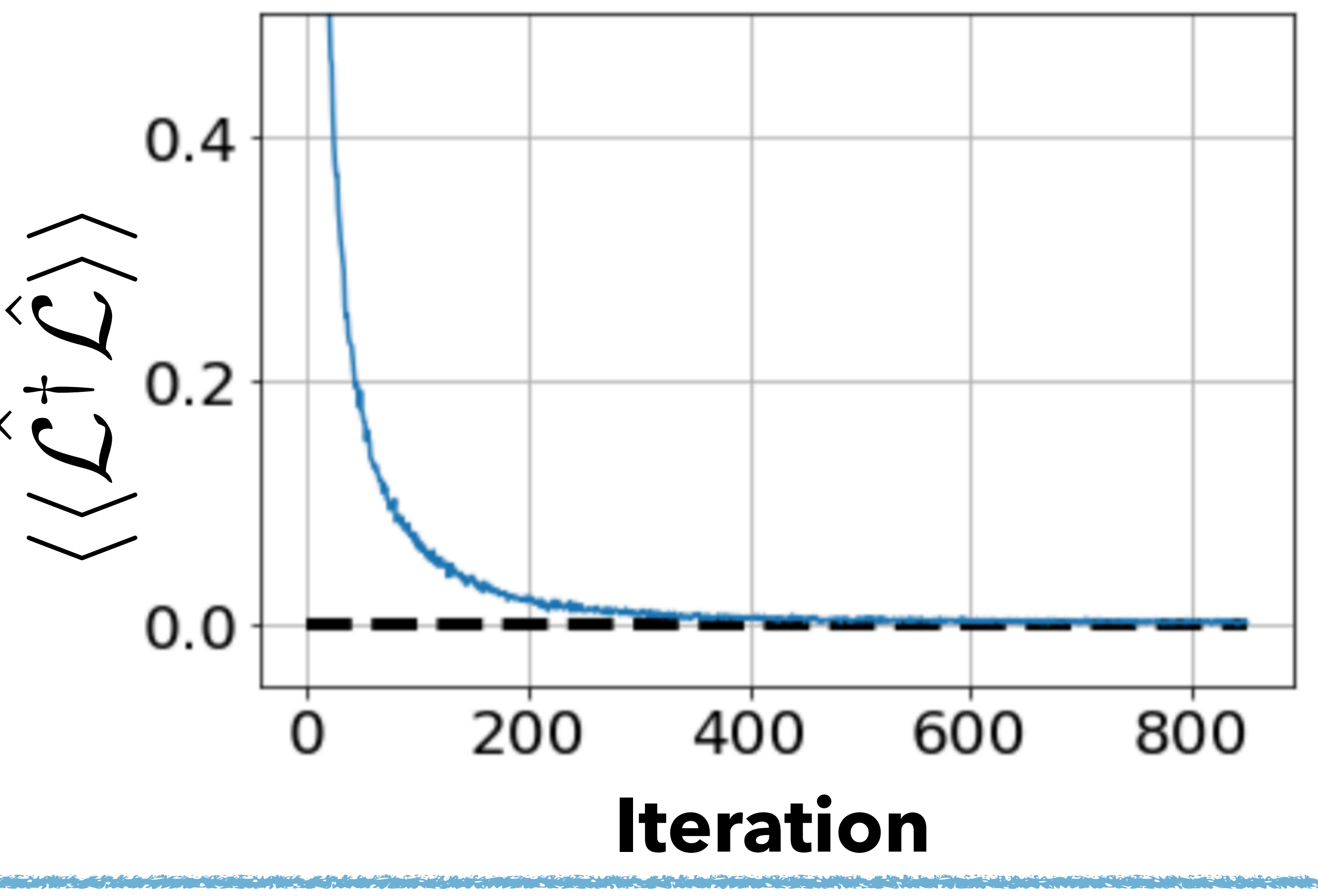
Step3

Ansatz selection

$$|\rho\rangle\rangle = \Big|$$  $$\Big\rangle\Big\rangle$$

Step4

Run optimization

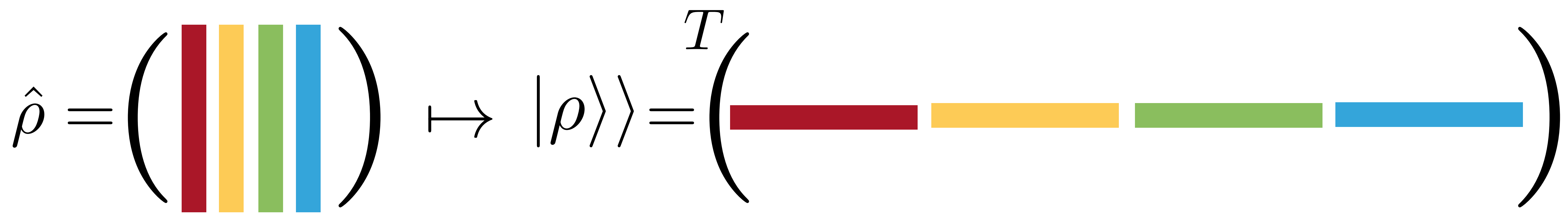


Precisely zero if exact

Step1: Vector representation

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

Under the representation of density matrix as pure state vector by *doubling the number of qubits.* Cui etal. PRL ('15)



$$\hat{\rho} = \sum_{\sigma\tau} \rho_{\sigma\tau} |\sigma\rangle\langle\tau| \quad \mapsto |\rho\rangle\rangle = \frac{1}{C} \sum_{\sigma\tau} \rho_{\sigma\tau} |\sigma, \tau\rangle\rangle$$

"physical" "fictitious"

Vector representation of Lindblad equation

Goal: Find $\hat{\mathcal{L}}|\hat{\rho}_{\mathrm{SS}}\rangle\rangle = 0$

$$\hat{\mathcal{L}}|\rho(t)\rangle\rangle = \left(-i \left(\hat{H} \otimes \hat{1} - \hat{1} \otimes \hat{H}^T \right) + \sum_i \gamma_i \hat{\mathcal{D}}[\hat{\Gamma}_i] \right) |\rho(t)\rangle\rangle$$

non anti-hermitian — Unitary dynamics — Dissipation, Non-unitary

where $\hat{\mathcal{D}}[\hat{\Gamma}_i] = \hat{\Gamma}_i \otimes \hat{\Gamma}_i^* - \frac{1}{2}\hat{\Gamma}_i^\dagger \hat{\Gamma}_i \otimes \hat{1} - \hat{1} \otimes \frac{1}{2}\hat{\Gamma}_i^T \hat{\Gamma}_i^*$ (e.g. $\hat{\Gamma}_i = \sigma_i^-$)

Step2: Steady state as "ground state"

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

Stationary state as zero eigenvector

Right eigenvectors of Lindblad operator $\hat{\mathcal{L}}$

$$\hat{\mathcal{L}}|\rho_n\rangle\rangle = \lambda_n|\rho_n\rangle\rangle \qquad e^{\hat{\mathcal{L}}t}|\rho_n\rangle\rangle = e^{\lambda_n t}|\rho_n\rangle\rangle$$

Eigenvalues satisfy

1. $\mathrm{Re}[\lambda_n] \leq 0$
2. At least one $\mathrm{Re}[\lambda_n] = 0$ (usually unique)

→ Steady state realized at $t \to \infty$

Idea: Ground-state search problem Cui etal. PRL ('15)

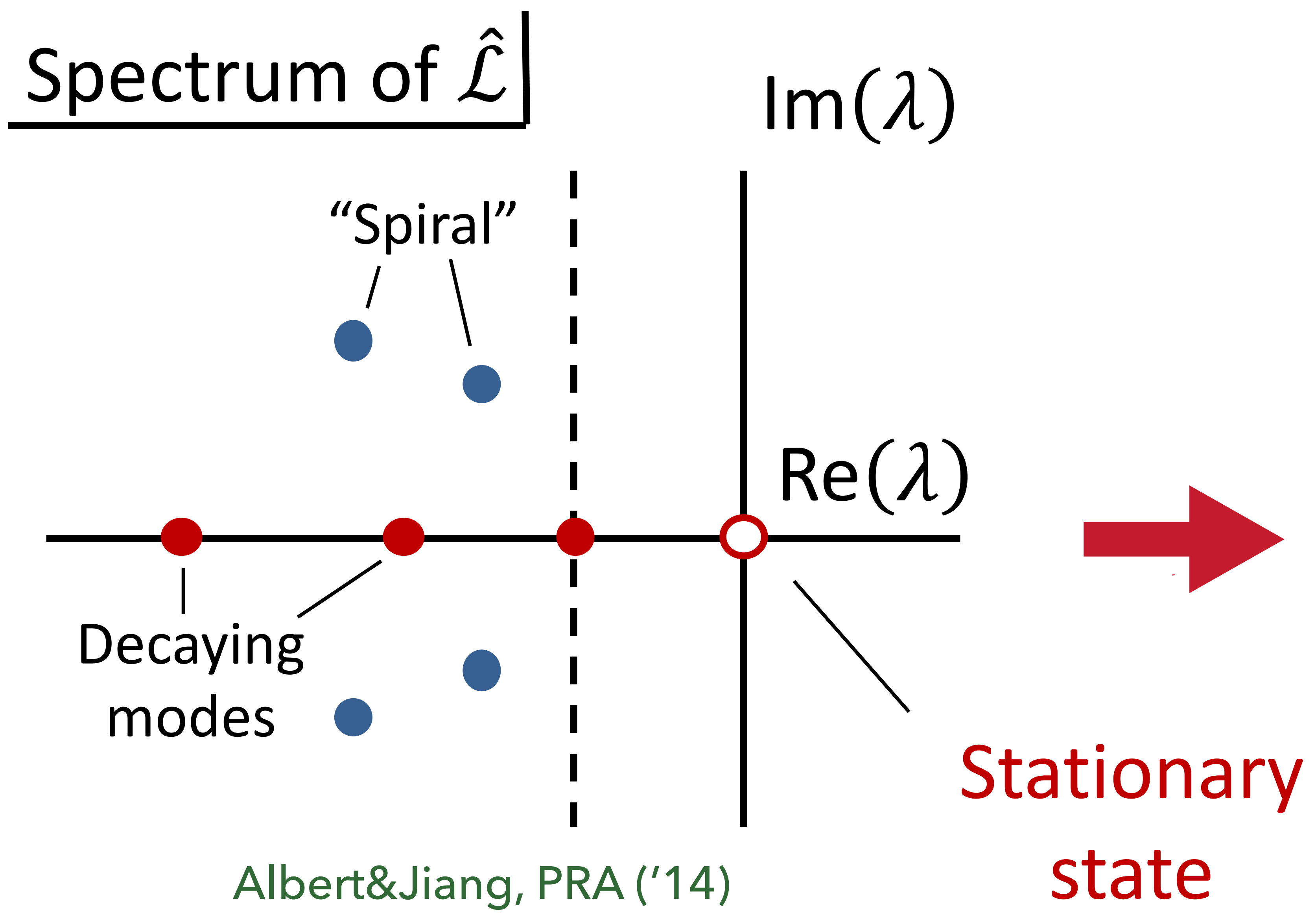
One can easily show that

$$\hat{\mathcal{L}}|\rho\rangle\rangle = 0 \longleftrightarrow \hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}|\rho\rangle\rangle = 0$$

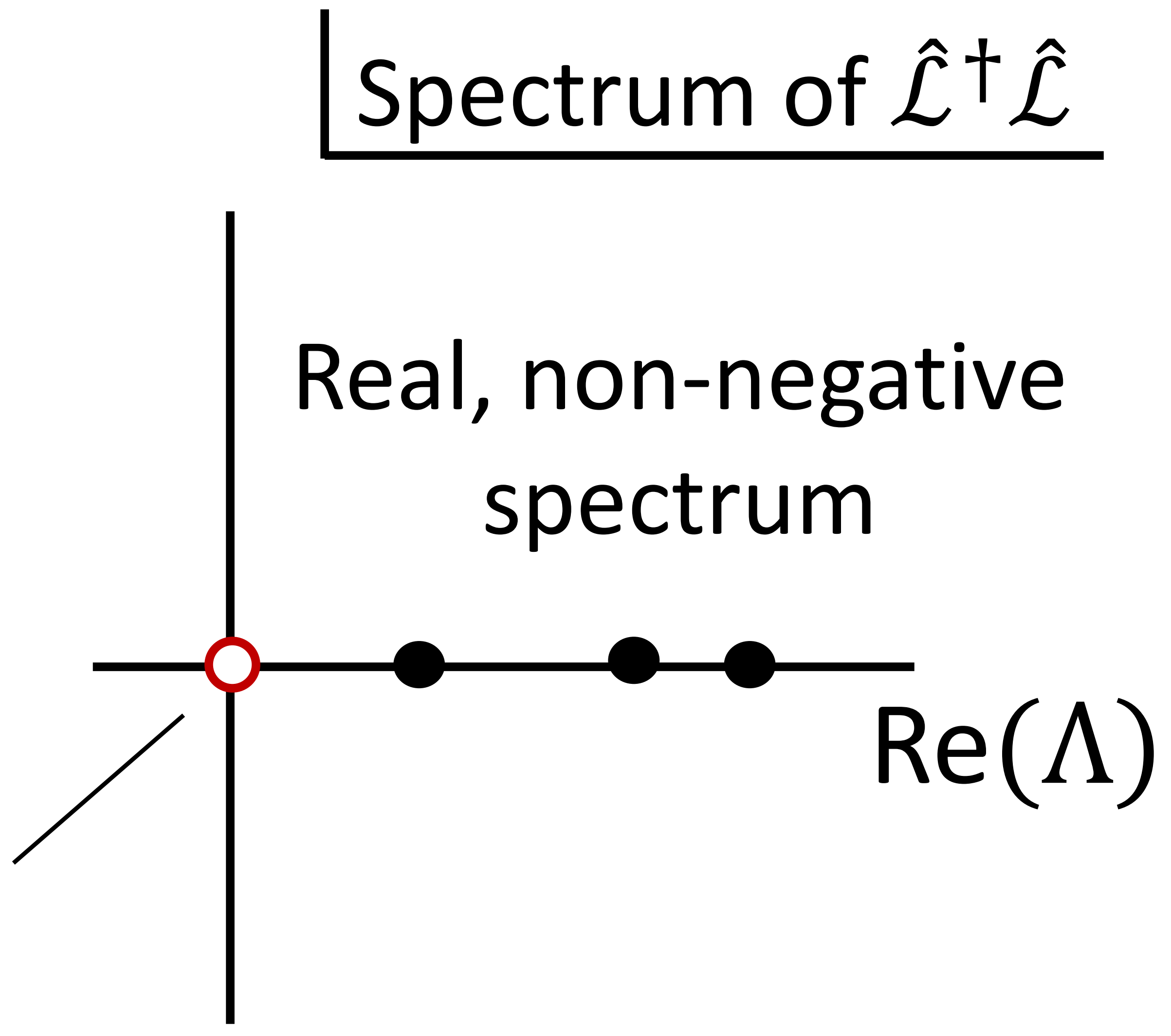
Steady state is obtained variationally by considering

$$0 = \underset{\rho}{\mathrm{argmin}} \frac{\langle\langle\rho|\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}|\rho\rangle\rangle}{\langle\langle\rho|\rho\rangle\rangle}$$

Our goal



Albert&Jiang, PRA ('14)



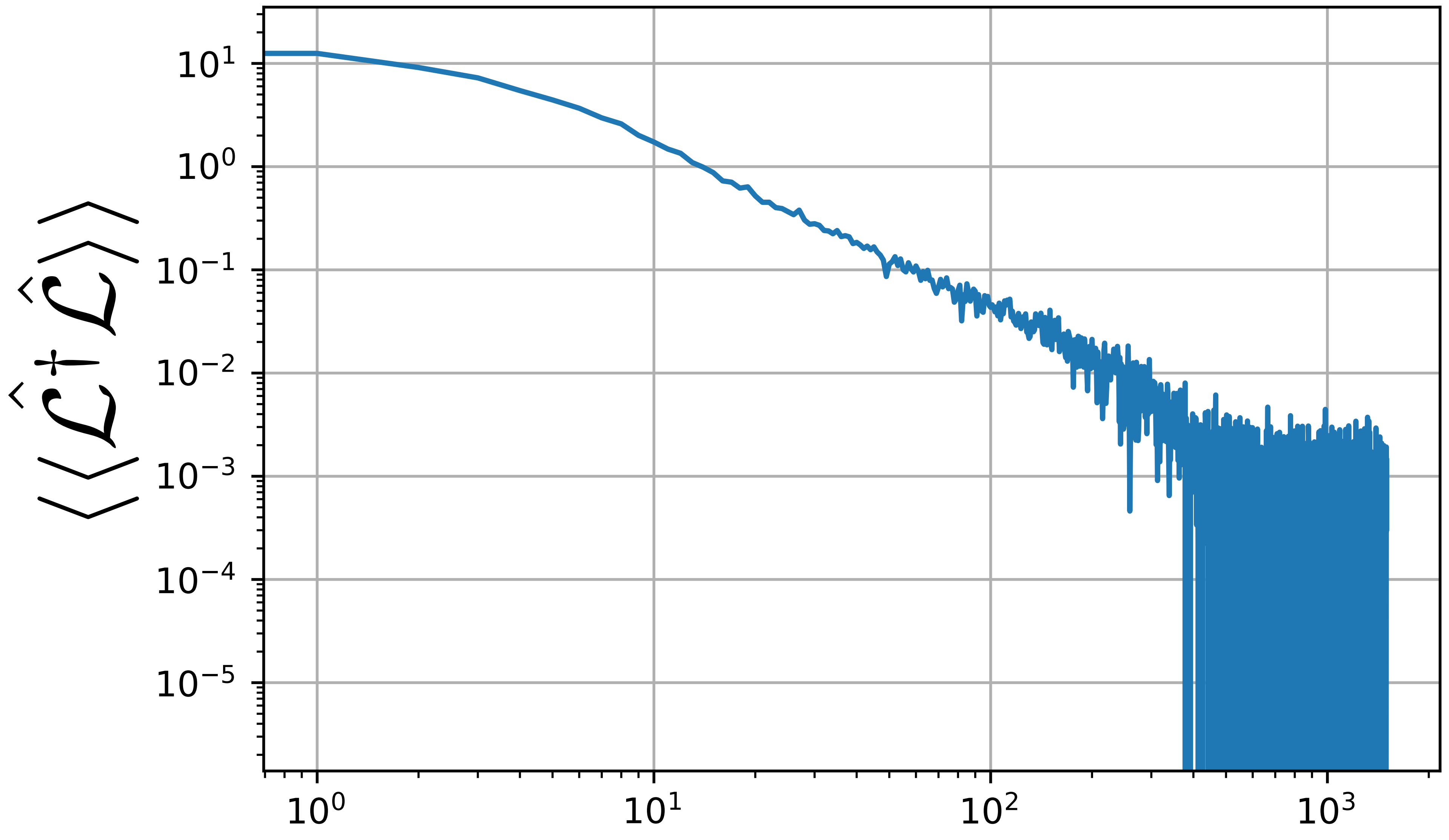
Step4: Run optimization

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

Optimization for 1d transverse-field Ising model (TFIM) under V=1, g=1, γ=1, Nspin = 8.

Convergence at $\langle\langle \hat{\mathcal{L}}^\dagger \hat{\mathcal{L}} \rangle\rangle < 10^{-3}$, **fidelity = 0.996** for α=4

Optimization with N_{it} = 1500, N_{samp} = 2000



Variational Monte Carlo Algorithm

0. Choose initial parameter p_0

1. Estimate parameter update δp

- Stochastic reconfiguration, i.e.,

  δp for imaginary time evolution by MC sampling

- $O(N_{samp}\alpha^2 N^2 + N_p^3)$ per update step

2. Update as $p \leftarrow p - \delta p$

3. Repeat 1-2 until convergence.

Model in above calculation : 1d TFIM + damping   Barreior etal. Nature ('11)   Carr etal. PRL ('13)

Realized in trapped ions, Rydberg atoms

$$\hat{H} = \frac{V}{4}\sum_{i=0}^{L-1} \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z + \frac{g}{2}\sum_{i=0}^{L-1} \hat{\sigma}_i^x,$$ with uniform damping as $\hat{\Gamma}_i = \hat{\sigma}_i^-$,

Results

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

e.g. Transverse-field Ising model w/ damping

Barreior etal. Nature ('11) Carr etal. PRL ('13)

Realized in trapped ions, Rydberg atoms

$$\hat{H} = \frac{V}{4}\sum_{i=0}^{L-1} \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z + \frac{g}{2}\sum_{i=0}^{L-1} \hat{\sigma}_i^x, \text{ with } \hat{\Gamma}_i = \hat{\sigma}_i^-,$$

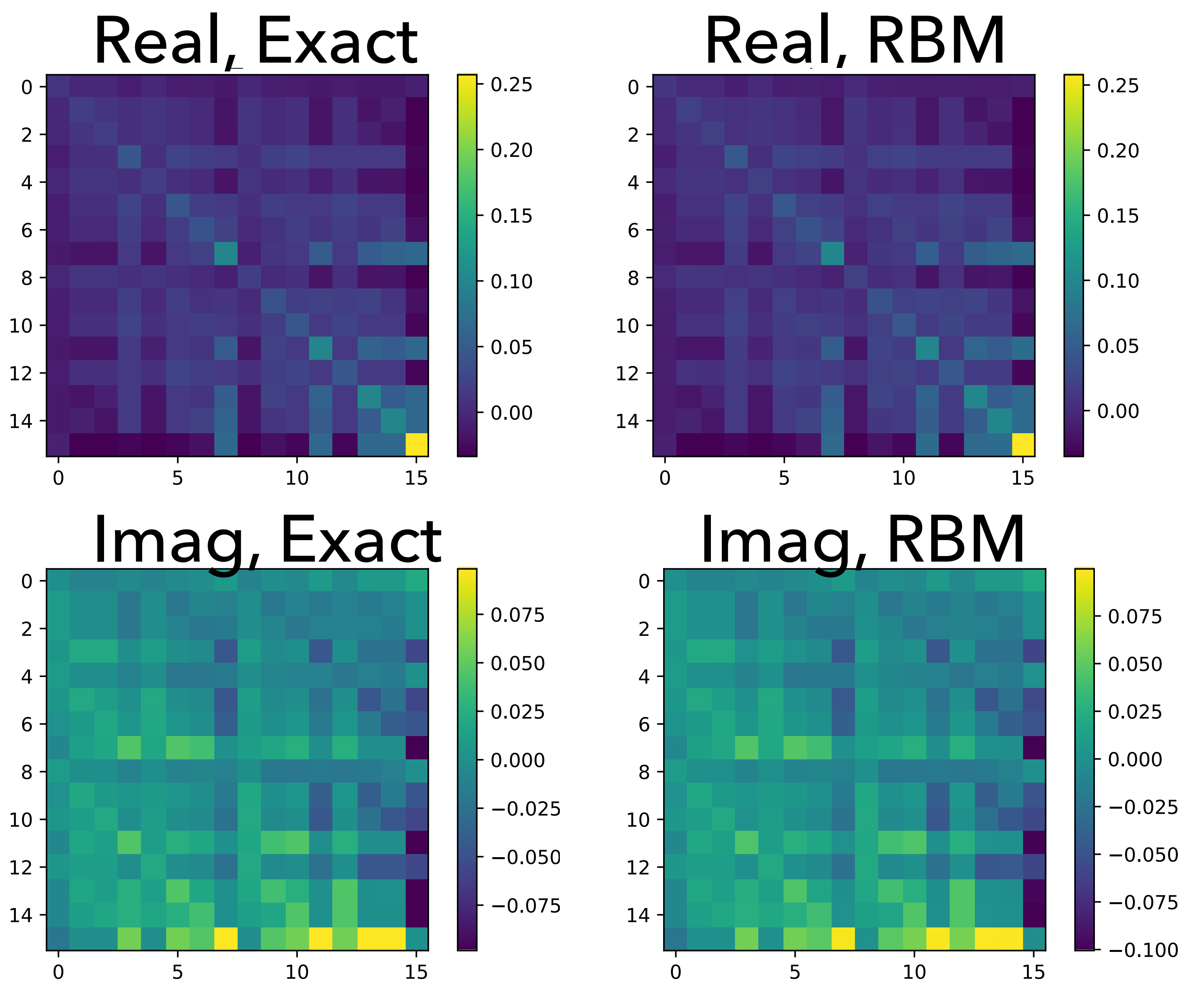
1d TFIM with damping

- Fidelity > 0.995 achieved up to L=16 (32 spins) at g/V = 0.3
- 40-fold #parameter reduction at strong field compared to MPS (L=16, reported by Hartmann&Carleo)

2d TFIM with damping

Jin etal. PRB ('18)

- Fidelity > 0.999 achieved for 2x2, 3x3 at g/V = 1
- Cost function optimized (~10^{-3}) up to 5x5



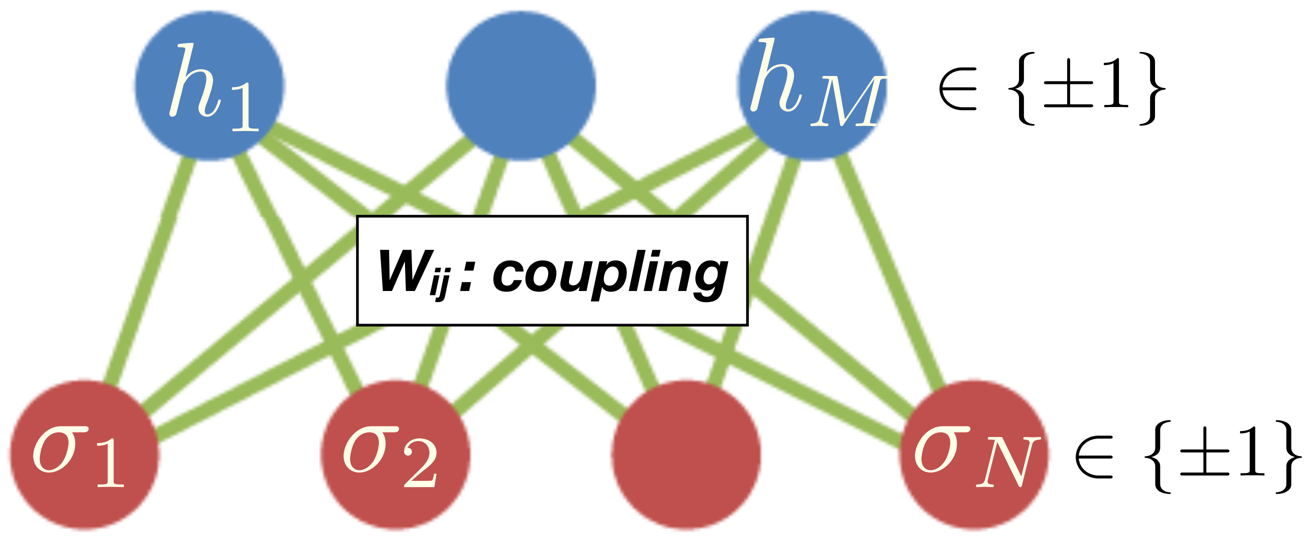
Real/Imaginary part of density matrices

2d TFIM, V=1, g=1, γ=1, fidelity>0.999

Summary

Introduction to neural quantum states

- Diverse application: GS, excited states, real-time evolution…
- Properties quite different from tensor nets
  e.g. quantum entanglement
- Mainly direct problem. What about inverse problem?



Solving dissipative many-body systems

Yoshioka&Hamazaki, Phys. Rev. B 99, 214306 (2019).

- Variational method for steady state in open quantum system demonstrated in
  - 1d/2d TFIM with damping
  - 1d XYZ with damping
- Both real-time and imaginary-time evolution possible

  Hartmann&Carleo PRL ('19)   Nagy&Savona PRL ('19)