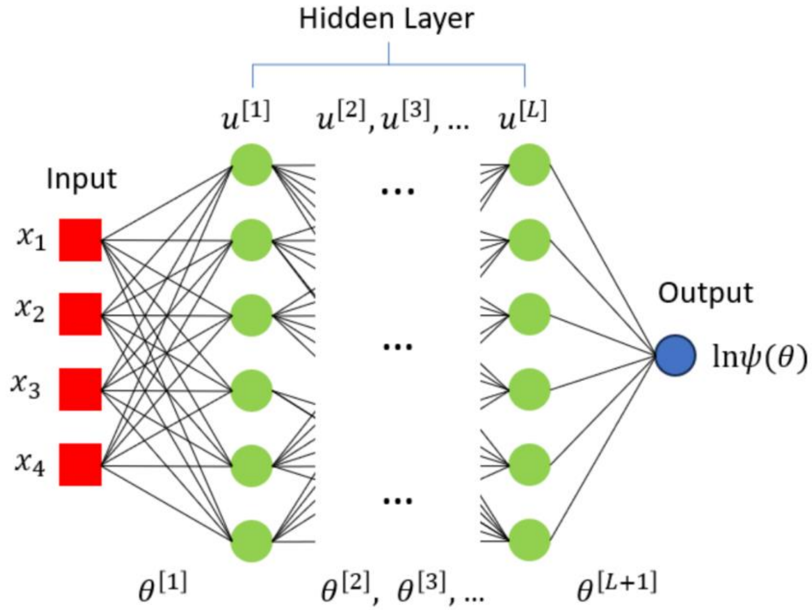
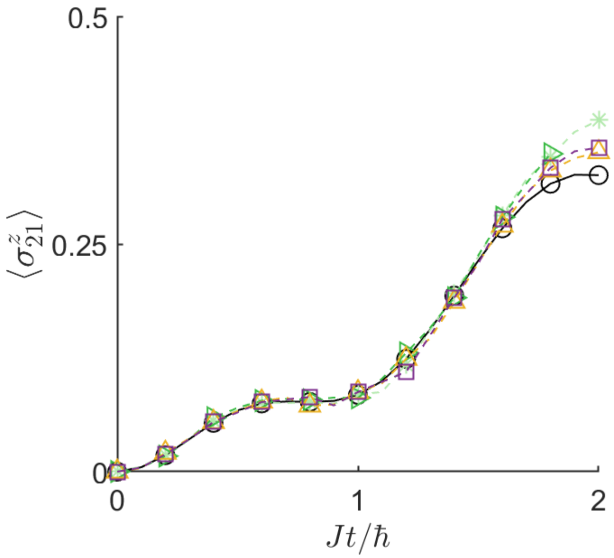
Paths towards time evolution with larger neural-network quantum states





Machine learning for quantum technology

Erlangen 2024

Dario Poletti

Singapore University of Technology and Design
Centre for Quantum Technologies
Majulab

Zhang et al., arxiv:2406.03381

QUANTUM ENGINEERING PROGRAMME SINGAPORE

Ministry of Education SINGAPORE

SG NSCC SINGAPORE National Supercomputing Centre

NATIONAL RESEARCH FOUNDATION
PRIME MINISTER'S OFFICE
SINGAPORE

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Centre for Quantum Technologies

MajuLab

Paths towards time evolution with larger neural-network quantum states

Principal Investigator
Dario Poletti

PostDocs
Noufal Jaseem
Wang Dingzu
Zhang Wenxuan

PhD Students
Rebecca Erbanni
Gauthameswar
Apimuk Sornsaeng

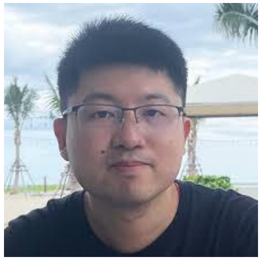
Collaborators on these topics

Guo Chu
Itai Arad
Lu Wei
Kavan Modi

Remmy Zen
Stephane Bressan
Christian Miniatura
Xu Xiansong
Xing Bo

Xu Xiansong

Xing Bo



Zhang Wenxuan

QUANTUM ENGINEERING PROGRAMME SINGAPORE

Ministry of Education SINGAPORE

SG NSCC SINGAPORE National Supercomputing Centre

NATIONAL RESEARCH FOUNDATION
PRIME MINISTER'S OFFICE
SINGAPORE

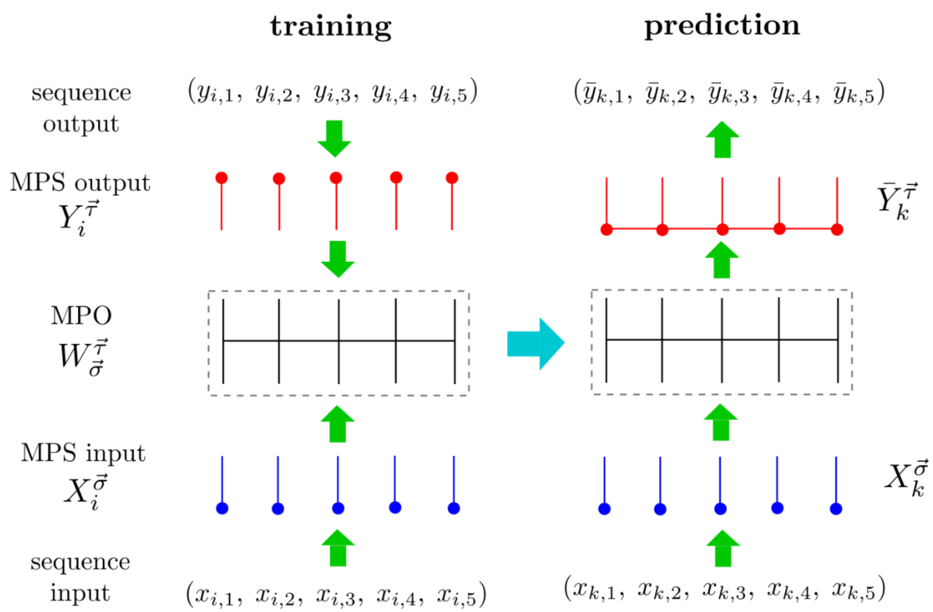
Zhang et al., arxiv:2406.03381

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

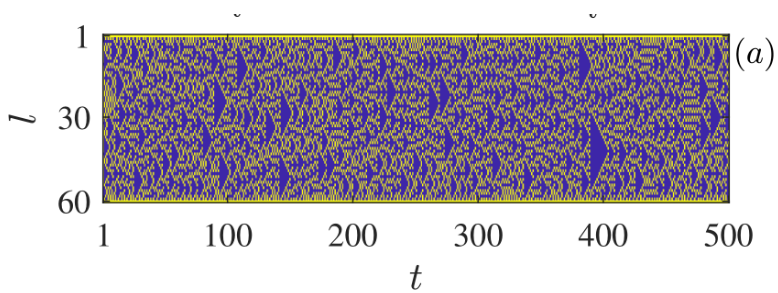
CQT Centre for Quantum Technologies

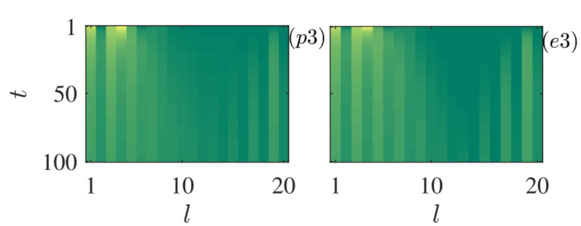
MajuLab

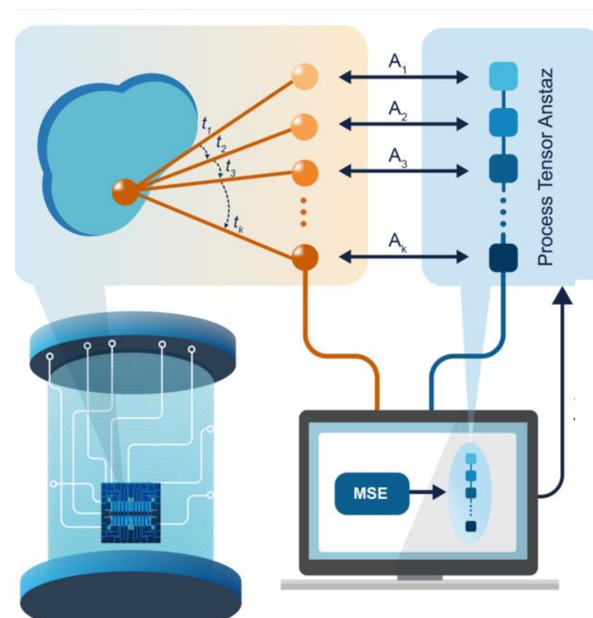
Using MPOs for sequence-to-sequence learning



Guo, et al., PRE 2018
C. Guo et al., PRA 2020
H.P. Casagrande et al., arXiv:2404.11768
X. Zhang, et al., arXiv:2312.06062
C. Guo, et al., manuscript in preparation





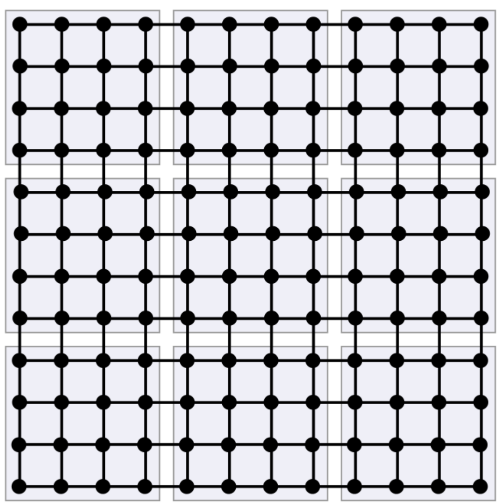


Block belief propagation and tensor networks

Merging tensor networks
with belief propagation

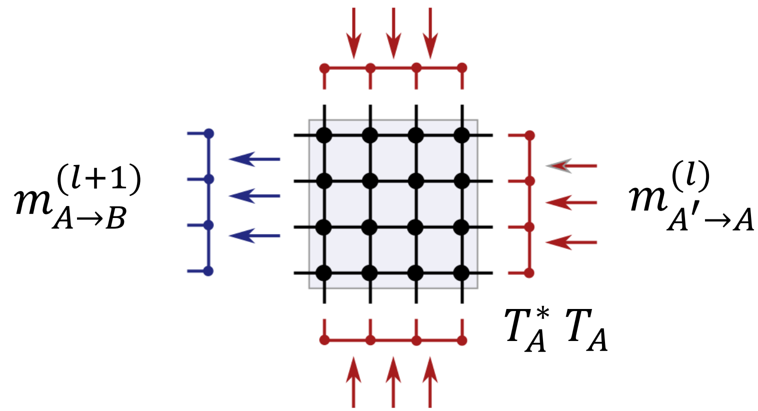
Divide the system into blocks

(a)



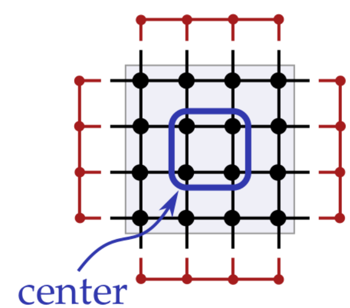
Update the message coming from one block

(b) blockBP iteration



Compute local quantities using messages from other blocks

(c) Local environment



Guo, D.P., Arad, PRB 2023

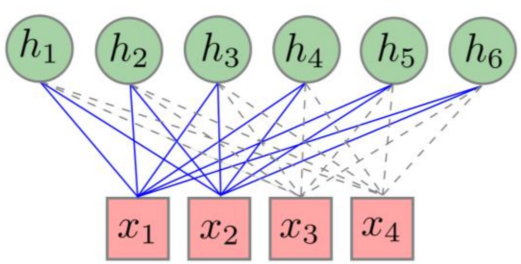
Plan of the presentation

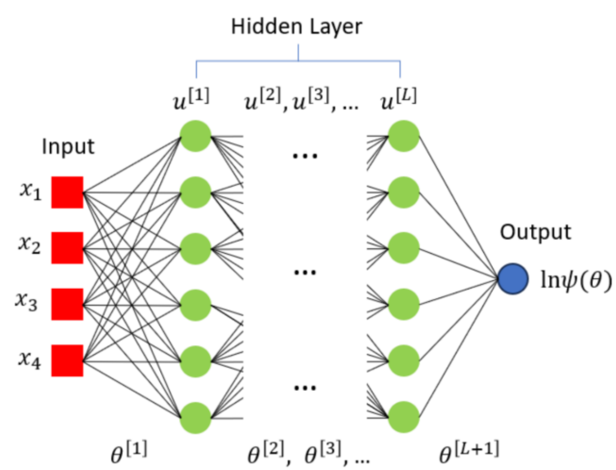
1) Time evolution from overlap
   - Loss function
   - Optimization approaches

2) Holomorphic neural networks

3) Quantum quench in the tilted Ising model

4) Conclusions and outlook





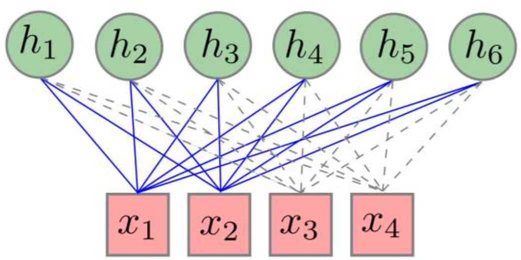
Plan of the presentation

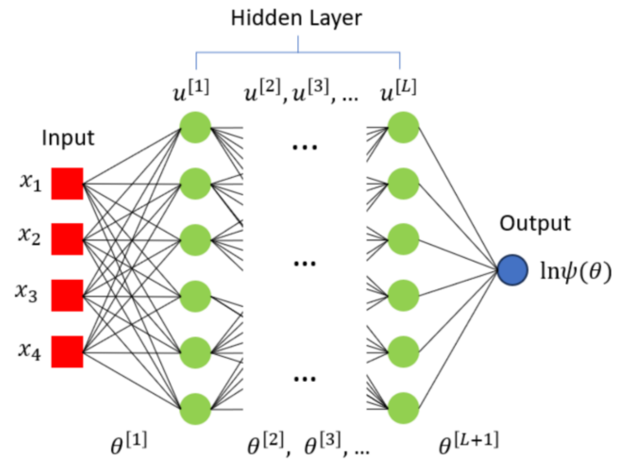
1) Time evolution from overlap
   - Loss function
   - Optimization approaches

2) Holomorphic neural networks

3) Quantum quench in the tilted Ising model

4) Conclusions and outlook





Time evolution with NQS

From Carleo and Troyer's 2017 paper we have been presented with one approach to simulate time evolution of many-body quantum systems with tensor networks

$$\frac{d\boldsymbol{\theta}(t)}{dt} = -\mathrm{i}\,\boldsymbol{S}^{-1}(t)\boldsymbol{F}(t)$$

where θ are the network parameters, and the equation is then derived from Schrodinger's equation. Also known as tVMC.

Here $\boldsymbol{S} = \boldsymbol{X}\boldsymbol{X}^\dagger$ and $\boldsymbol{F} = \boldsymbol{X}\boldsymbol{f}$ with

$$X_{m,k} = \frac{1}{\sqrt{N_s}}\left(\left(\frac{\partial \ln[\psi(\boldsymbol{x}_k;\boldsymbol{\theta})]}{\partial\theta_m}\right)^* - \left\langle\frac{\partial \ln[\psi(\boldsymbol{x};\boldsymbol{\theta})]}{\partial\theta_m}\right\rangle^*\right)$$

and

$$f_k = \frac{1}{\sqrt{N_s}}(E_{loc}(\boldsymbol{x}_k) - \langle E_{loc}(\boldsymbol{x})\rangle)$$

Carleo and Troyer, Science 2017
Schmitt and Heyl, Phys. Rev. Lett. 125,
Lee, Patil, Zhang, and Hsieh, Phys. Rev. Res. 3, 023095 (2021)
Schmitt et al., Sci. Adv. 8, 6850 (2022)

Time evolution with NQS

A different approach aims to find directly the neural network which best approximates the evolved one. Also known as p-tVMC.

For instance,

$$\left\|\psi[\theta_{n+1}] - \Phi^{\Delta t}\left(\psi[\theta_n]\right)\right\|$$

Gutierrez and Mendl, Quantum 2022

Donatella, Denis, Le Boite', Ciuti, Phys. Rev. A 108, 022210 (2023)

Sinibaldi, Giuliani, Carleo, Vicentini, Quantum 7, 1131 (2023)

Gravina, Savona, Vicentini, arxiv:2410.10720

Time evolution with NQS

A different approach aims to find directly the neural network which best approximates the evolved one. Also known as p-tVMC.

For instance,

$$\left\|\psi[\theta_{n+1}] - \Phi^{\Delta t}\left(\psi[\theta_n]\right)\right\|$$

This can be done in various ways, either expanding the unitary evolution

$$\psi[\theta_{n+1}] \approx \psi[\theta_n] - i\Delta t H\left(\frac{\psi[\theta_{n+1}] + \psi[\theta_n]}{2}\right)$$

Gutierrez and Mendl, Quantum 2022
Donatella, Denis, Le Boite', Ciuti, Phys. Rev. A 108, 022210 (2023)
Sinibaldi, Giuliani, Carleo, Vicentini, Quantum 7, 1131 (2023)
Gravina, Savona, Vicentini, arxiv:2410.10720

Time evolution with NQS

A different approach aims to find directly the neural network which best approximates the evolved one. Also known as p-tVMC.

For instance,

$$\left\|\psi[\theta_{n+1}] - \Phi^{\Delta t}\left(\psi[\theta_n]\right)\right\|$$

This can be done in various ways, either expanding the unitary evolution

$$\psi[\theta_{n+1}] \approx \psi[\theta_n] - i\Delta t H\left(\frac{\psi[\theta_{n+1}] + \psi[\theta_n]}{2}\right)$$

or not

$$\mathcal{I}_{\mathrm{loc}}(\sigma,\eta) = 1 - \frac{\langle\sigma|\mathcal{U}\,|\Psi_\theta\rangle}{\langle\sigma|\Psi_{\bar{\theta}}\rangle}\frac{\langle\eta|\mathcal{U}^\dagger\,|\Psi_{\bar{\theta}}\rangle}{\langle\eta|\Psi_\theta\rangle}$$

Gutierrez and Mendl, Quantum 2022
Donatella, Denis, Le Boite', Ciuti, Phys. Rev. A 108, 022210 (2023)
Sinibaldi, Giuliani, Carleo, Vicentini, Quantum 7, 1131 (2023)
Gravina, Savona, Vicentini, arxiv:2410.10720

Neural network quantum states: **time evolution**

Cost function

$$C^{U}_{\psi_t,\psi_{t'}} = \frac{\langle\psi_{t'}|U|\psi_t\rangle\langle\psi_t|U^\dagger|\psi_{t'}\rangle}{\langle\psi_{t'}|\psi_{t'}\rangle\langle\psi_t|\psi_t\rangle}$$

Neural network quantum states: **time evolution**

Cost function

Forward evolution

$$C^{U}_{\psi_t,\psi_{t'}} = \frac{\langle\psi_{t'}|U|\psi_t\rangle\langle\psi_t|U^\dagger|\psi_{t'}\rangle}{\langle\psi_{t'}|\psi_{t'}\rangle\langle\psi_t|\psi_t\rangle}$$

Backward evolution

Neural network quantum states: **time evolution**

Cost function

An effective Hamiltonian

$$C^{U}_{\psi_t,\psi_{t'}} = \frac{\langle \psi_{t'} | U | \psi_t \rangle \langle \psi_t | U^{\dagger} | \psi_{t'} \rangle}{\langle \psi_{t'} | \psi_{t'} \rangle \langle \psi_t | \psi_t \rangle}$$

Neural network quantum states: **time evolution**

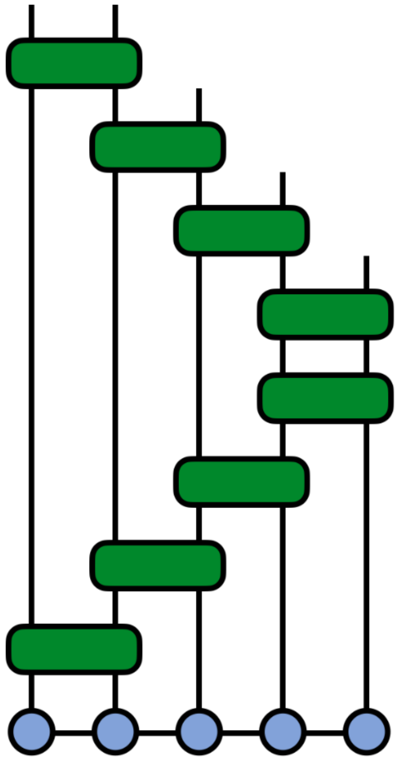
Cost function

An effective Hamiltonian

$$C^{U}_{\psi_t,\psi_{t'}} = \frac{\langle\psi_{t'}|U|\psi_t\rangle\langle\psi_t|U^\dagger|\psi_{t'}\rangle}{\langle\psi_{t'}|\psi_{t'}\rangle\langle\psi_t|\psi_t\rangle}$$

$$U(dt) = \prod_{i=1}^{L-1} U_{2,i}(dt/2) \prod_{j=L-1}^{1} U_{2,j}(dt/2)$$

Suzuki-Trotter decomposition



In this way the U used at each step is sparse

Neural network quantum states: **approaches to train larger networks**

Cost function

$$
\begin{aligned}
C^{U}_{\psi_t,\psi_{t'}} &= \sum_{\boldsymbol{x},\boldsymbol{y}} P_{\psi_{t'}}(\boldsymbol{x}) P_{\psi_t}(\boldsymbol{y}) E^{U}_{\psi_t \psi_{t'}}(\boldsymbol{x}) E^{U}_{\psi_{t'} \psi_t}(\boldsymbol{y}) \\
&= \sum_{\boldsymbol{x}} P_{\psi_{t'}}(\boldsymbol{x}) E^{U}_{loc}(\boldsymbol{x}),
\end{aligned}
$$

this is readily amenable to use Stochastic Reconfiguration for the time evolution.

$$\boldsymbol{\theta} = \boldsymbol{\theta} - \gamma \mathbf{S}^{-1}\mathbf{F}$$

Carleo and Troyer, Science 2017 Sorella et al. Jour. Chem. Phys. 2007

Neural network quantum states: **approaches to train larger networks**

Cost function

$$C^{U}_{\psi_t,\psi_{t'}} = \sum_{\boldsymbol{x},\boldsymbol{y}} P_{\psi_{t'}}(\boldsymbol{x}) P_{\psi_t}(\boldsymbol{y}) E^{U}_{\psi_t \psi_{t'}}(\boldsymbol{x}) E^{U}_{\psi_{t'} \psi_t}(\boldsymbol{y})$$
$$= \sum_{\boldsymbol{x}} P_{\psi_{t'}}(\boldsymbol{x}) E^{U}_{loc}(\boldsymbol{x}),$$

this is readily amenable to use Stochastic Reconfiguration for the time evolution.

$$\boldsymbol{\theta} = \boldsymbol{\theta} - \gamma \mathbf{S}^{-1}\mathbf{F}$$

But inverting the S matrix becomes more and more demanding

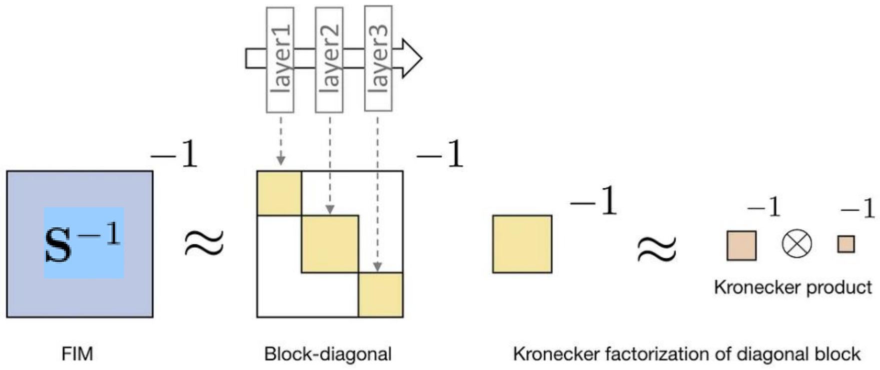
$$S_{m,n} = \left\langle \left(\frac{\partial \ln[\psi(\boldsymbol{x};\boldsymbol{\theta})]}{\partial \theta_m} \right)^{*} \frac{\partial \ln[\psi(\boldsymbol{x};\boldsymbol{\theta})]}{\partial \theta_n} \right\rangle - \left\langle \frac{\partial \ln[\psi(\boldsymbol{x};\boldsymbol{\theta})]}{\partial \theta_m} \right\rangle^{*} \left\langle \frac{\partial \ln[\psi(\boldsymbol{x};\boldsymbol{\theta})]}{\partial \theta_n} \right\rangle$$

Matrix of size
Number of parameters × Number of parameters

Carleo and Troyer, Science 2017 Sorella et al. Jour. Chem. Phys. 2007

Neural network quantum states: **approaches to train larger networks**

(1) K-FAC approximates the inverse by turning it into a block diagonal.

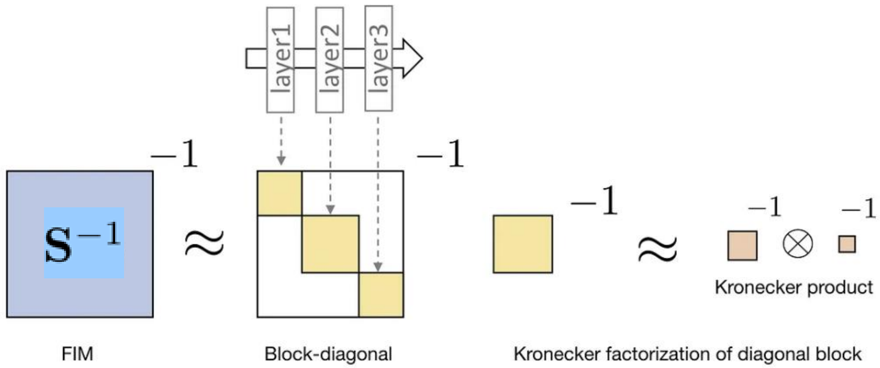


J. Martens and R. Grosse, Proc. Mach. Lear. Res. 37,2408 (2015)

From towardsdatascience.com

Neural network quantum states: **approaches to train larger networks**

(1) K-FAC approximates the inverse by turning it into a block diagonal.



J. Martens and R. Grosse, Proc. Mach. Lear. Res. 37,2408 (2015)

From towardsdatascience.com

(2) minSR turns the matrix to invert from being limited by the number of parameters to be limited by the number of samples which can be much smaller than the number of parameters.

$$(\boldsymbol{S} + \lambda \mathbb{1}_p)^{-1} \boldsymbol{F}^U = (\boldsymbol{X}\boldsymbol{X}^\dagger + \lambda \mathbb{1}_p)^{-1} \boldsymbol{X} \boldsymbol{f}_E^U$$
$$= \boldsymbol{X}(\boldsymbol{X}^\dagger \boldsymbol{X} + \lambda \mathbb{1}_s)^{-1} \boldsymbol{f}_E^U$$

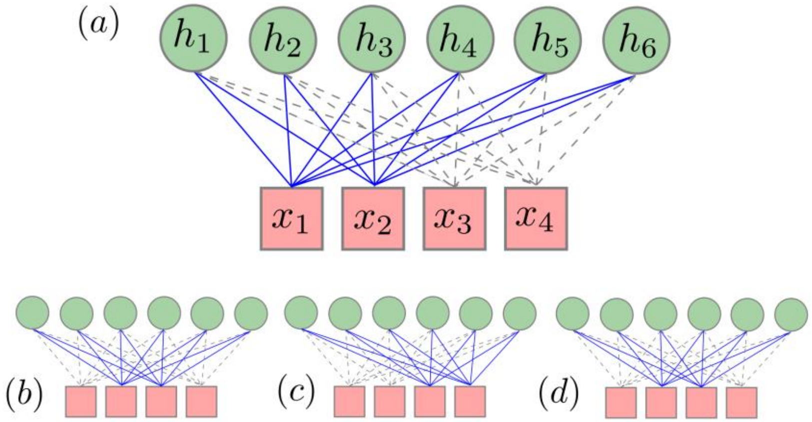
Matrix of size

Number of samples × Number of samples

Chen and Heyl, Nature Physics 2024
Rende, et al., Comm. Physics 2024

Neural network quantum states: **approaches to train larger networks**

(3) Sequential Local Optimization SLO is a DMRG-inspired approach which we have developed for ground state search



where the key is to optimize only a portion of the network, but to also ensure **overlap between consecutively optimized portions**.

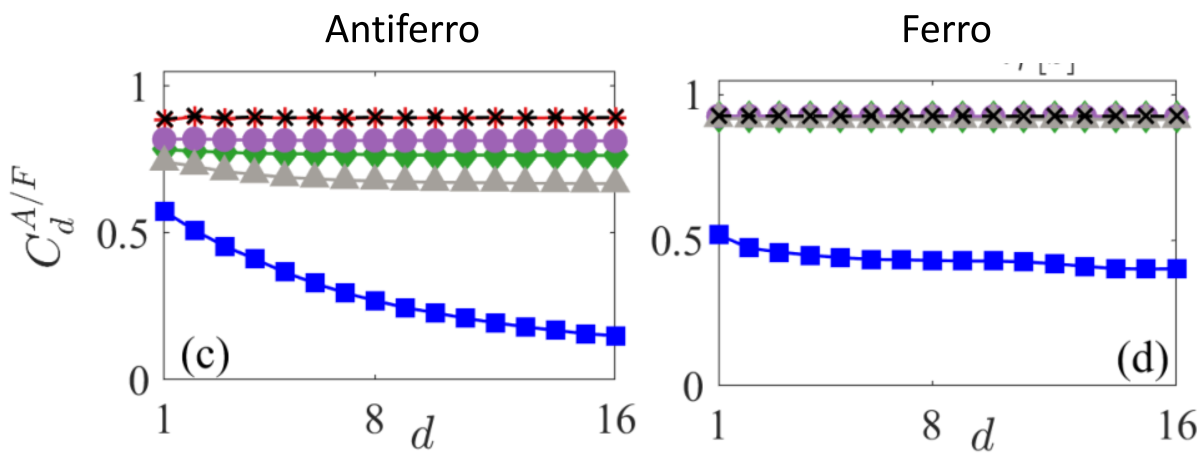
Zhang et al, PRB 2023

Neural network quantum states: **approaches to train larger networks**

For example, we consider an Ising model whose ground state can be Ferromagnetic, Antiferromagnetic or Paramagnetic.

$$H = \sum_{\langle i,j \rangle} J \sigma_i^z \sigma_j^z - \sum_i (h_z \sigma_i^z + h_x \sigma_i^x)$$

Antiferro and Ferro correlations after 400s

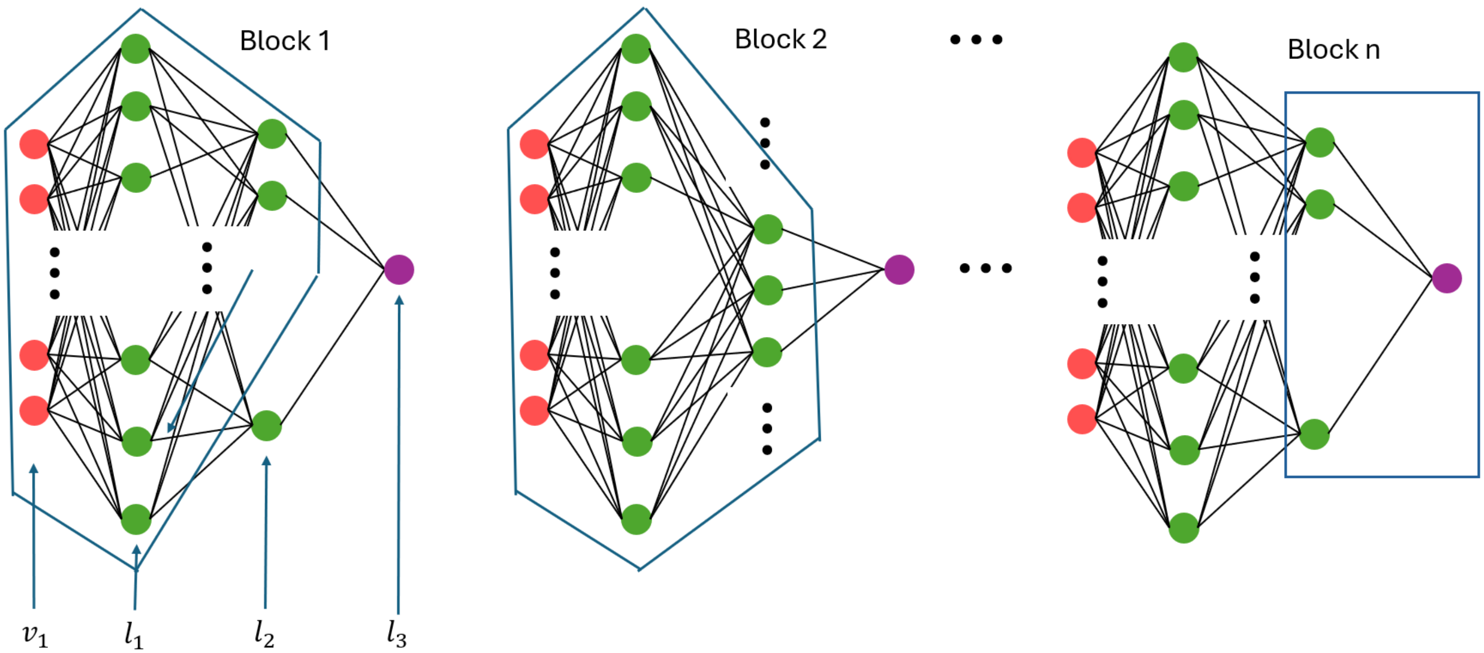


L = 32

The symbol $*$ represents 2-site block, ◆ 4-site block, ● 8-site block, ▲ 16-site block and ■ represents updating all parameters.

Neural network quantum states: **approaches to train larger networks**

For the time evolution, with our FFNN, we use this type of SLO which resulted to be more stable.



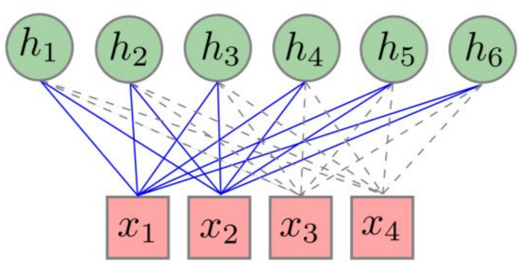
Plan of the presentation

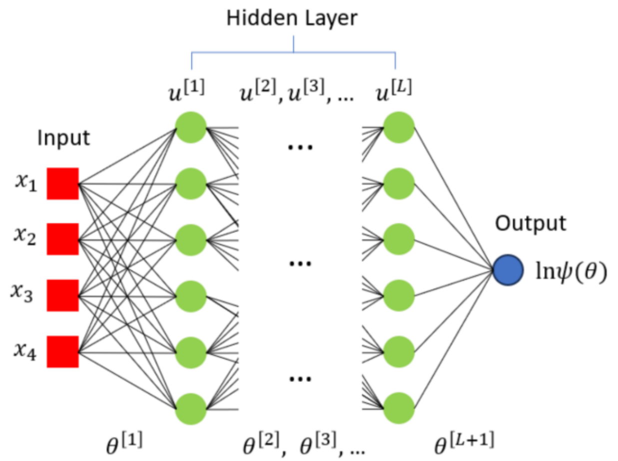
1) Time evolution from overlap
   - Loss function
   - Optimization approaches

2) Holomorphic neural networks

3) Quantum quench in the tilted Ising model

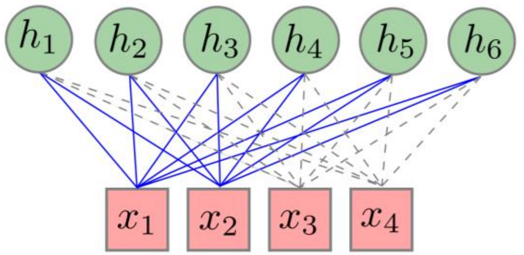
4) Conclusions and outlook





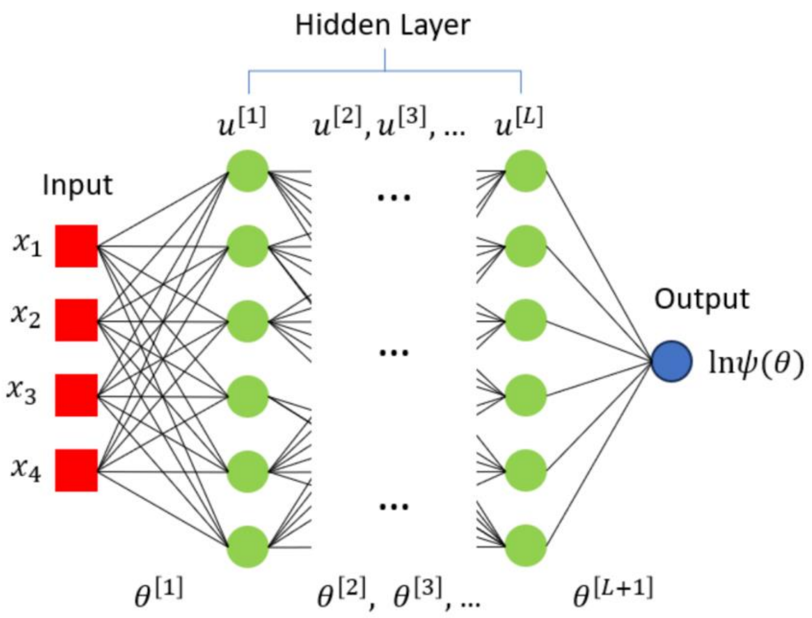
Neural network quantum states considered and their properties

Restricted Boltzmann machines



$$\psi(\boldsymbol{x};\boldsymbol{\theta})_{\mathrm{RBM}} = \exp\left(\sum_l a_l x_l\right)\prod_{h=1}^{H} 2\cosh\left(b_h + \sum_{l=1}^{L} W_{h,l}x_l\right)$$

Feed-forward neural networks



$$\boldsymbol{u}^{[k]} = \boldsymbol{W}^{[k]} f(\boldsymbol{u}^{[k-1]}) + \boldsymbol{b}^{[k]}$$

$$f(\boldsymbol{u}) = \boldsymbol{u} - \frac{1}{3}\boldsymbol{u}^3 + \frac{2}{15}\boldsymbol{u}^5$$

$$\psi(\boldsymbol{x};\boldsymbol{\theta})_{\mathrm{FNN}} = \exp\left(\boldsymbol{u}^{[K]}\right)$$

Neural network quantum states considered and their properties

Our main object of interest is the loss function

$$\mathcal{F}(|\psi_\theta\rangle, |\phi\rangle) = \frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}$$

Neural network quantum states considered and their properties

Our main object of interest is the loss function

$$\mathcal{F}(|\psi_\theta\rangle, |\phi\rangle) = \frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}$$

Evolved past wave function

Neural network quantum states considered and their properties

Our main object of interest is the loss function

$$\mathcal{F}(|\psi_\theta\rangle, |\phi\rangle) = \frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}$$

$$\frac{\partial}{\partial z^*}(f^* f) = \left(\frac{\partial}{\partial z^*} f^*\right) f$$

$$\partial_{\boldsymbol{\theta}^*}\mathcal{F} = \frac{\partial}{\partial \boldsymbol{\theta}^*}\left(\frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}\right)$$

$$= \frac{\frac{\partial\langle\psi|}{\partial\theta^*}|\phi\rangle}{\langle\psi|\psi\rangle} \cdot \frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle} - \frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle} \cdot \frac{1}{(\langle\psi|\psi\rangle)^2} \cdot \frac{\partial\langle\psi|}{\partial\boldsymbol{\theta}^*}|\psi\rangle$$

$$\partial_{\boldsymbol{\theta}^*}\mathcal{F} = \mathbb{E}_{\boldsymbol{y}\sim\pi_\phi}\left[\frac{\psi(\boldsymbol{y})}{\phi(\boldsymbol{y})}\right] \cdot \mathbb{E}_{\boldsymbol{x}\sim\pi_\psi}\left[(\Delta J(\boldsymbol{x}))^* \cdot \frac{\phi(\boldsymbol{x})}{\psi(\boldsymbol{x})}\right]$$

$$\Delta J(\boldsymbol{x}) = J(\boldsymbol{x}) - \mathbb{E}_{\boldsymbol{z}\sim\pi_\psi}\left[J(\boldsymbol{z})\right]$$

$$J(\boldsymbol{x}) = \partial_{\boldsymbol{\theta}} \log\psi(\boldsymbol{x})$$

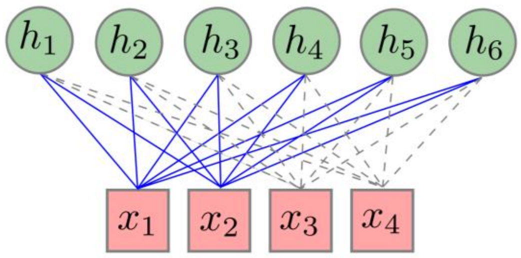
Plan of the presentation

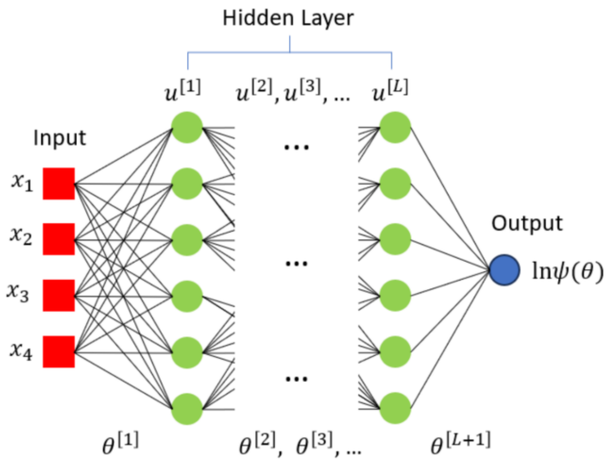
1) Time evolution from overlap
   - Loss function
   - Optimization approaches

2) Holomorphic neural networks

3) Quantum quench in the tilted Ising model

4) Conclusions and outlook





Quantum quench in the tilted Ising model

Tilted Ising model (TIM)

$$H = J \sum_{<i,j>} \sigma_i^z \sigma_j^z - h_x \sum_k \sigma_k^x - h_z \sum_m \sigma_m^z$$

Hamiltonian is divided into blocks:

$$H = \sum_p H_p$$

with

$$H_p = J\sigma_p^z \sigma_q^z - h_x(n_p \sigma_p^x + n_q \sigma_q^x) - h_z(n_p \sigma_p^z + n_q \sigma_q^z)$$

n_α with $\alpha = p, q$ is coordination number of the site and for 1D TIM

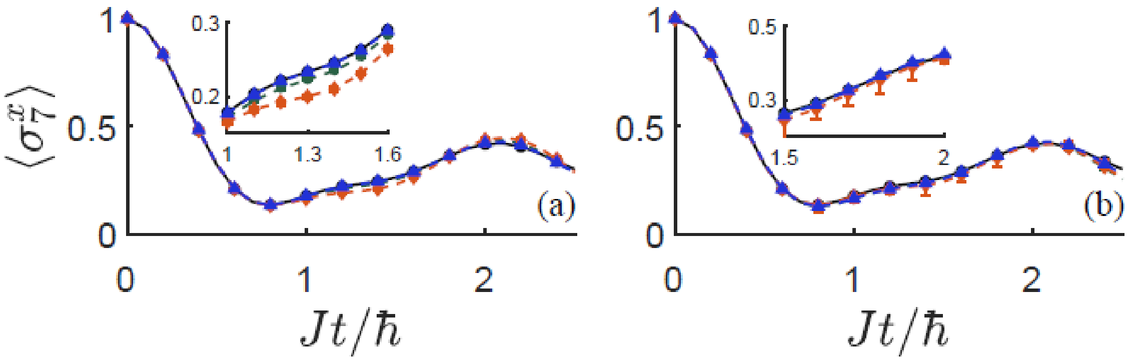
$$\begin{cases} n_\alpha = 1, \; \alpha \in \{1, L\} \\ n_\alpha = \frac{1}{2}, \; \alpha \notin \{1, L\} \end{cases}$$

The unitary operator can include more than one local operator

$$U = \exp(-i \sum_{\{p\}} H_p \cdot dt/2)$$

Quantum quench in the tilted Ising model

FFNN [14,56,42,1] RBM α=5

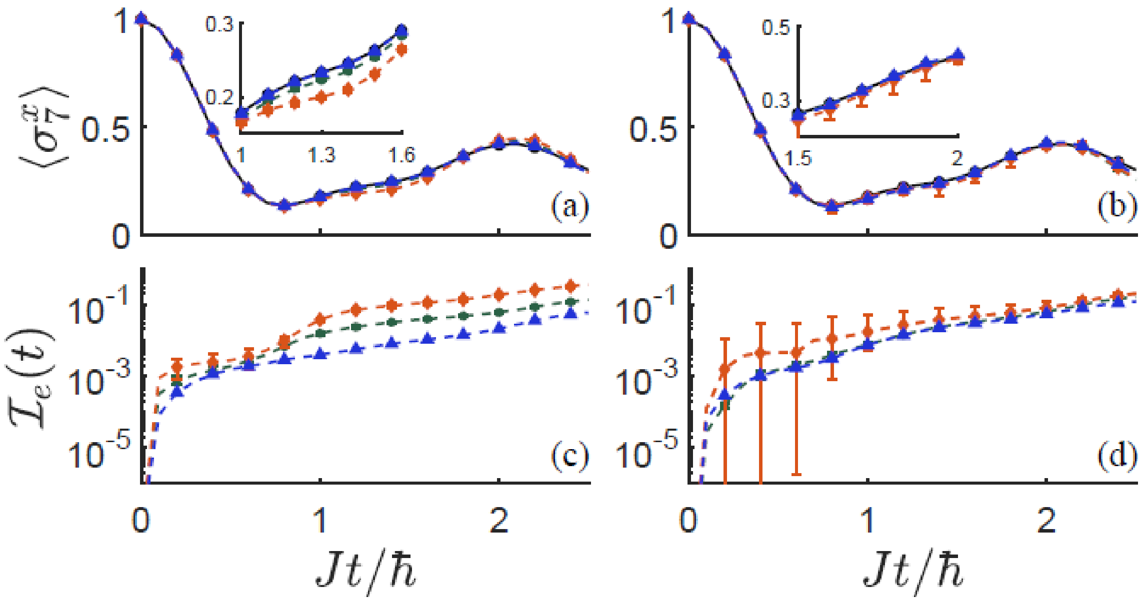


L = 14

● Exact Diagonalization
◊ Runge-Kutta (dt=10^{-3})
* Overlap (dt=10^{-1})
Δ Overlap Full summation (dt=10^{-1})

Quantum quench in the tilted Ising model

FFNN [14,56,42,1]　　　RBM α=5



Cumulated error from exact state

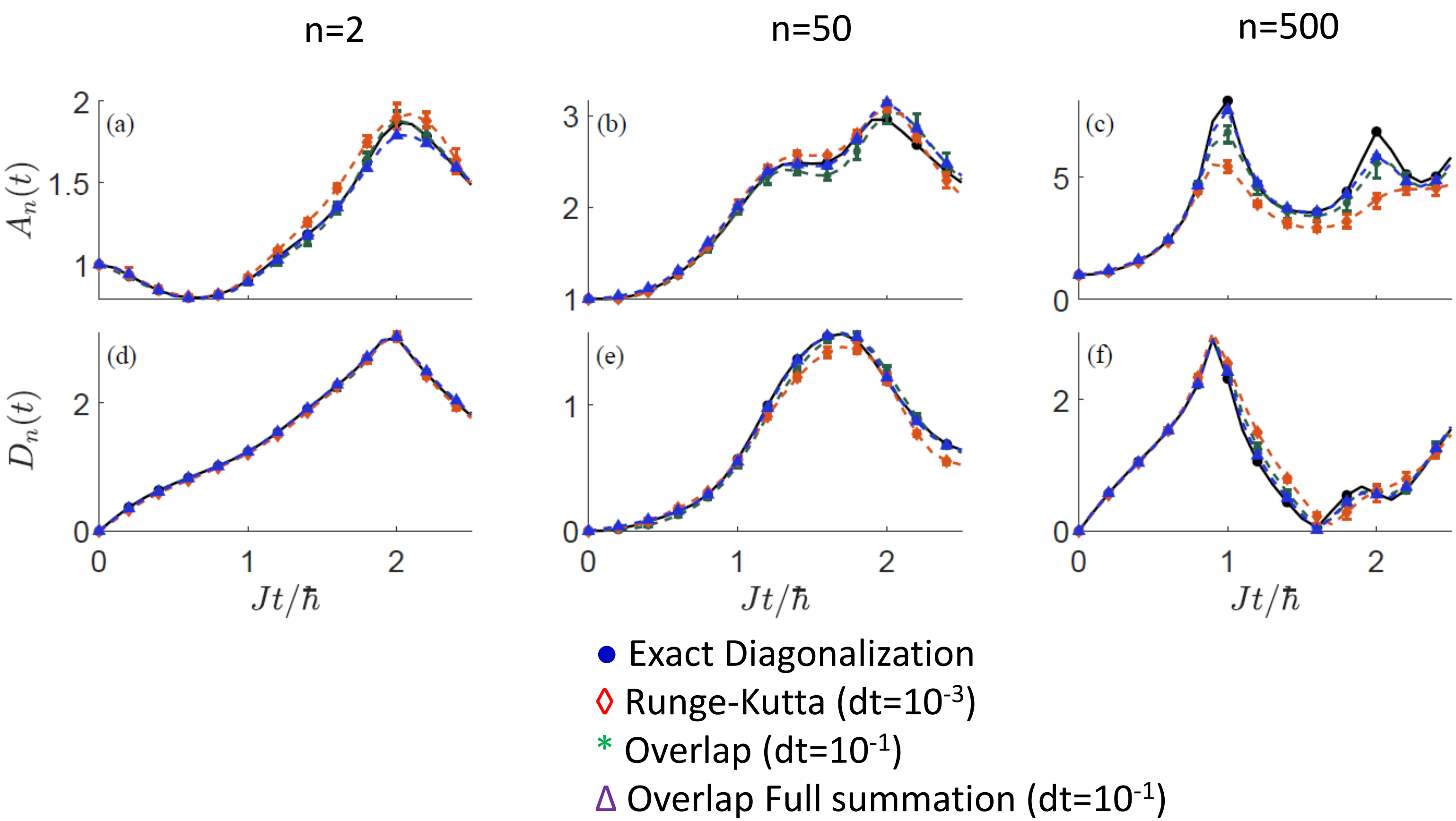
$$I_e(t) = \left| 1 - \frac{\langle \psi_t | \psi_{e,t} \rangle \langle \psi_{e,t} | \psi_t \rangle}{\langle \psi_t | \psi_t \rangle \langle \psi_{e,t} | \psi_{e,t} \rangle} \right|$$

$$\mathcal{I}_e(t) = \int_0^t I_e(t) dt$$

L = 14

● Exact Diagonalization
◊ Runge-Kutta (dt=10^{-3})
* Overlap (dt=10^{-1})
Δ Overlap Full summation (dt=10^{-1})

Quantum quench in the tilted Ising model



Ratio of amplitude between most likely and n-th likely configuration

$$A_n(t) = \left| \frac{\psi_t(\boldsymbol{x}_M)}{\psi_t(\boldsymbol{y}_n)} \right|$$

Relative phase between most likely and n-th likely configuration

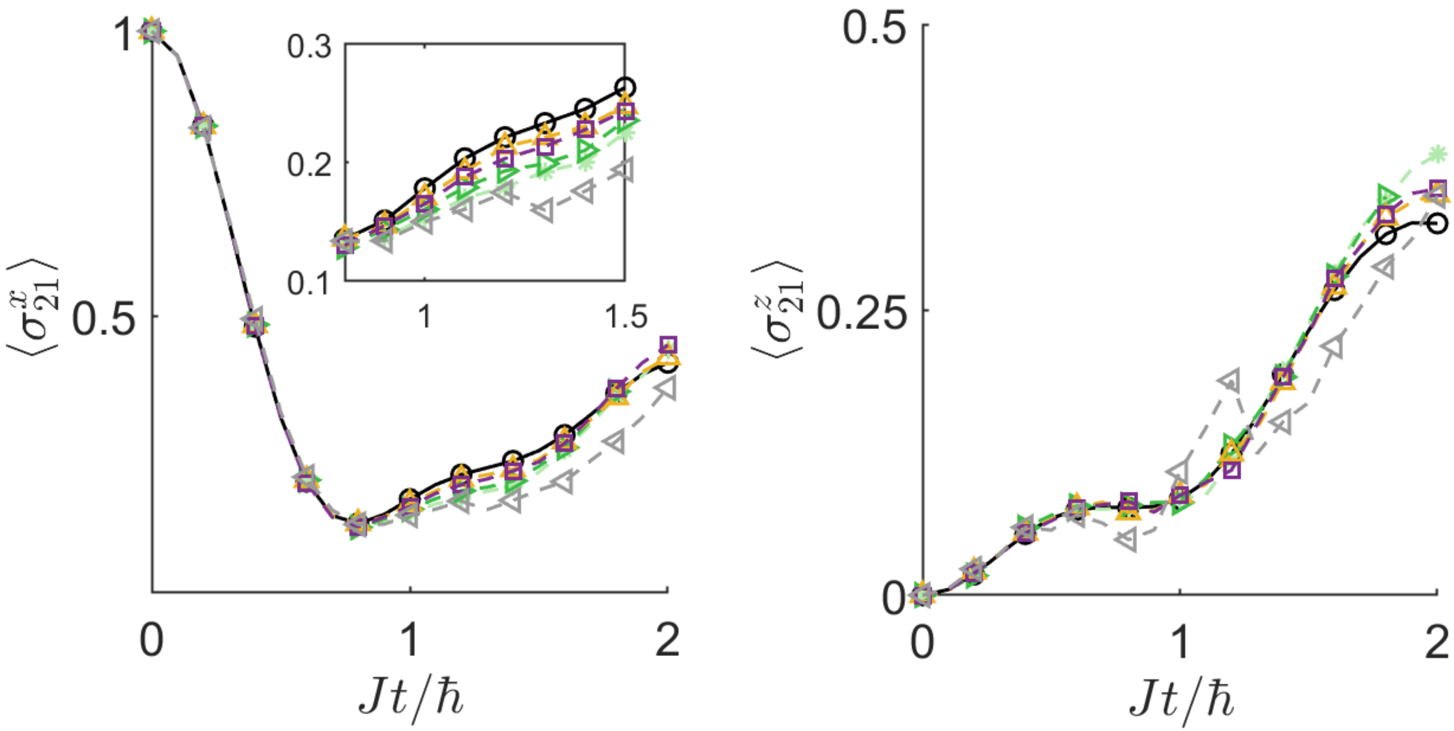
$$D_n(t) = \min(d_n(t), 2\pi - d_n(t))$$

$$d_n(t) = |\operatorname{Arg}(\psi_t(\boldsymbol{x}_M)) - \operatorname{Arg}(\psi_t(\boldsymbol{y}_n))|$$

Quantum quench in the tilted Ising model

L = 40

FFNN [L, 4L, 3L, 1]



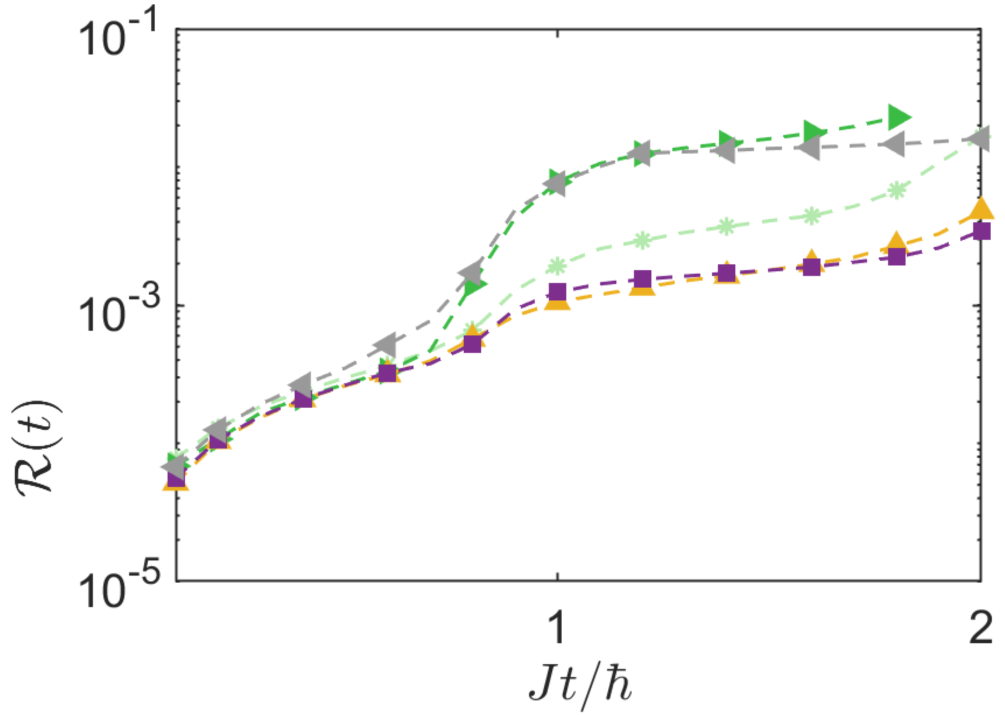
- ● t-MPS
- * K-FAC with 10^4 samples
- ▷ K-FAC with 10^5 samples
- □ SLO with 10^4 samples
- Δ minSR with 10^4 samples
- ◁ AdamW with 10^4 samples

- minSR seems to perform better
- SLO is more accurate than K-FAC
- K-FAC does not seem to improve significantly with more sample

Quantum quench in the tilted Ising model

L = 40

FFNN [L, 4L, 3L, 1]



Cumulated error from state we have

$$\mathcal{R}(t) = \int_0^t \sum_l I^{U_{d,l}}_{\psi_t, \psi_t'} dt$$

- ● t-MPS
- * K-FAC with 10^5 samples
- ▷ K-FAC with 10^4 samples
- □ SLO with 10^4 samples
- Δ minSR with 10^4 samples
- ◁ AdamW with 10^4 samples

- minSR seems to perform better
- SLO is more accurate than K-FAC
- K-FAC does not seem to improve significantly with more sample

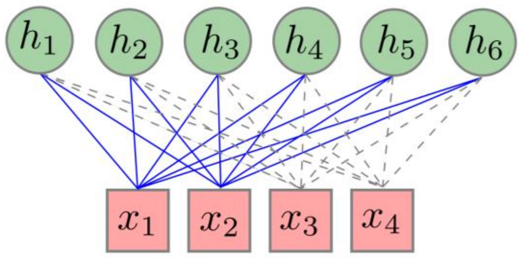
Plan of the presentation

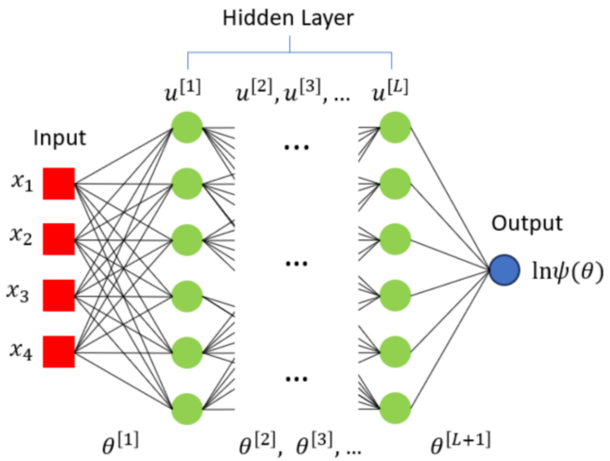
1) Time evolution from overlap
   - Loss function
   - Optimization approaches

2) Holomorphic neural networks

3) Quantum quench in the tilted Ising model

4) Conclusions and outlook





Conclusions and outlook

Ground state search-like methods can be used for time evolution, like minSR, K-FAC and SLO.

minSR seems, in our limited exploration, the better performing and more stable one.

Comparable performance can be achieved by SLO.

QUANTUM ENGINEERING PROGRAMME SINGAPORE

Ministry of Education SINGAPORE

National Supercomputing Centre SINGAPORE

NATIONAL RESEARCH FOUNDATION PRIME MINISTER'S OFFICE SINGAPORE

Zhang et al., arxiv:2406.03381

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Centre for Quantum Technologies

MajuLab

Conclusions and outlook

Ground state search-like methods can be used for time evolution, like minSR, K-FAC and SLO.

minSR seems, in our limited exploration, the better performing and more stable one.

Comparable performance can be achieved by SLO.

Outlook:

- Translate this to "lighter" and more expressive networks
- Further understand the role of sampling and how control variate can help (see latest work by Gravina, Savona and Vicentini)

QUANTUM ENGINEERING PROGRAMME SINGAPORE

Ministry of Education SINGAPORE

SG NSCC SINGAPORE National Supercomputing Centre

NATIONAL RESEARCH FOUNDATION PRIME MINISTER'S OFFICE SINGAPORE

Zhang et al., arxiv:2406.03381

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

CQT Centre for Quantum Technologies

MajuLab

Thank you!

Principal Investigator
Dario Poletti

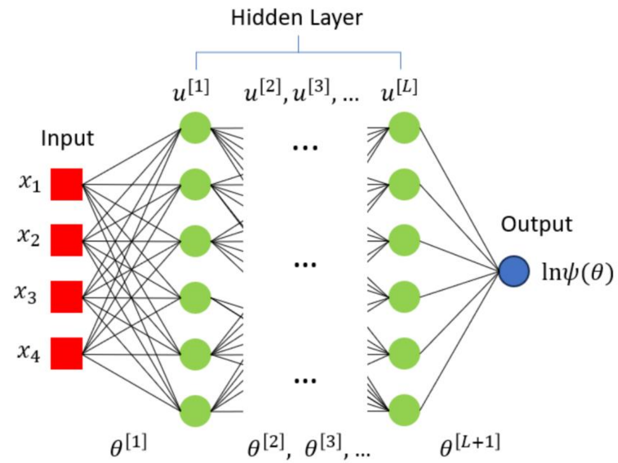
PostDocs
Noufal Jaseem
Wang Dingzu
Zhang Wenxuan

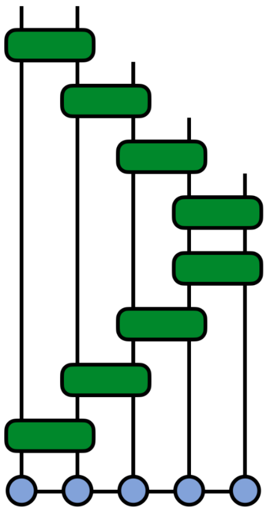
PhD Students
Rebecca Erbanni
Gauthameswar
Apimuk Sornsaeng

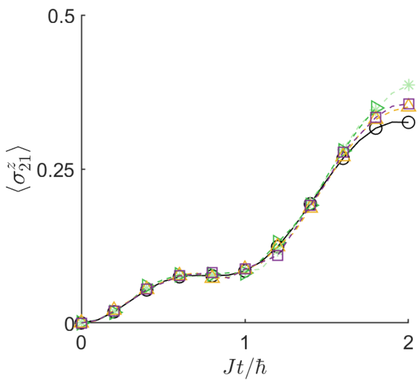
Collaborators on these topics

Guo Chu
Itai Arad
Lu Wei
Kavan Modi

Remmy Zen
Stephane Bressan
Christian Miniatura
Xu Xiansong
Xing Bo







QUANTUM ENGINEERING PROGRAMME SINGAPORE

Ministry of Education SINGAPORE

SG NSCC SINGAPORE National Supercomputing Centre

NATIONAL RESEARCH FOUNDATION
PRIME MINISTER'S OFFICE
SINGAPORE

Zhang et al., arxiv:2406.03381

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Centre for Quantum Technologies

MajuLab