Information scrambling and the learning landscape of a quantum machine learning algorithm

Sabre Kais
Department of Electrical and Computer Engineering, Chemistry and physics

NC State University
https://ece.ncsu.edu/people/skais/

Quantum Machine-Learning

Goal

Develop quantum algorithms
for
quantum simulations on quantum devices

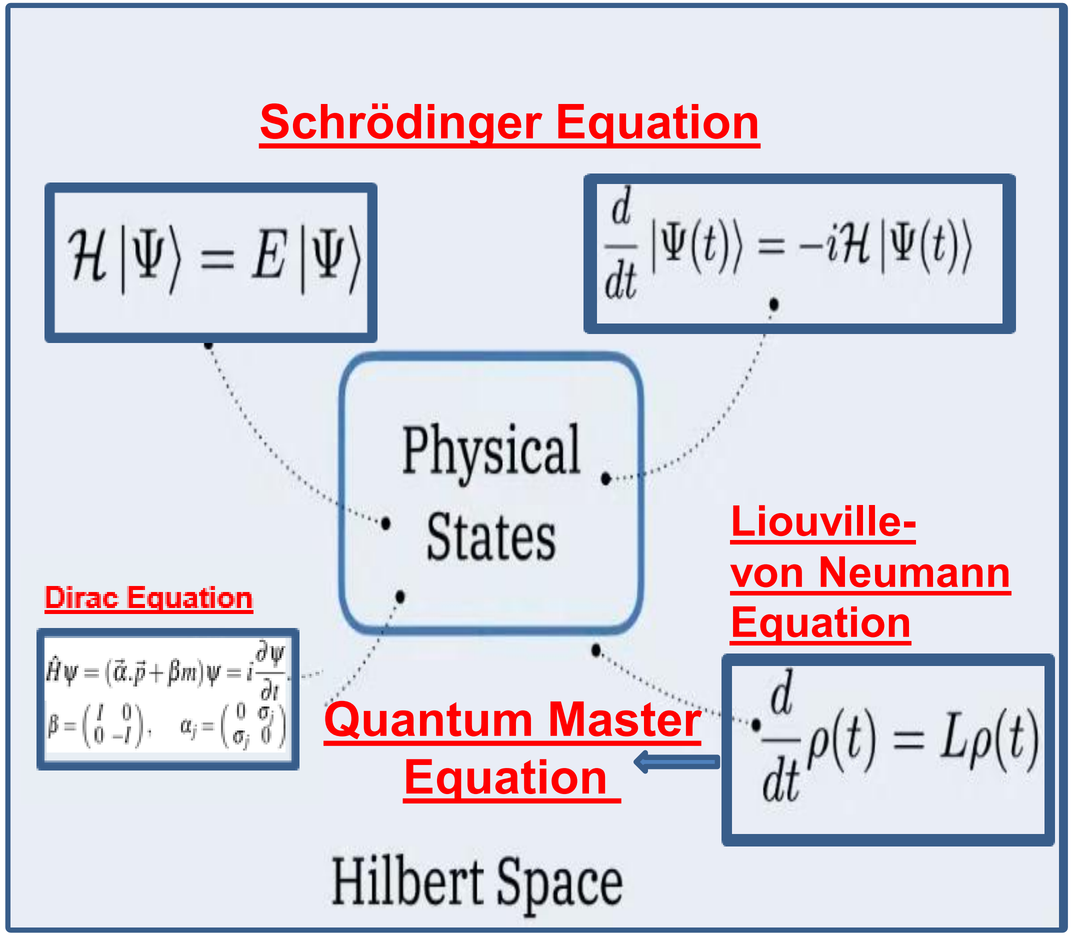
Complex Many-Body Systems

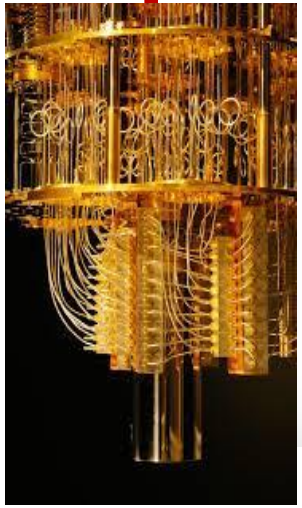
- **Electronic Structure**
- **Quantum Dynamics**

Complex Many-Body Systems

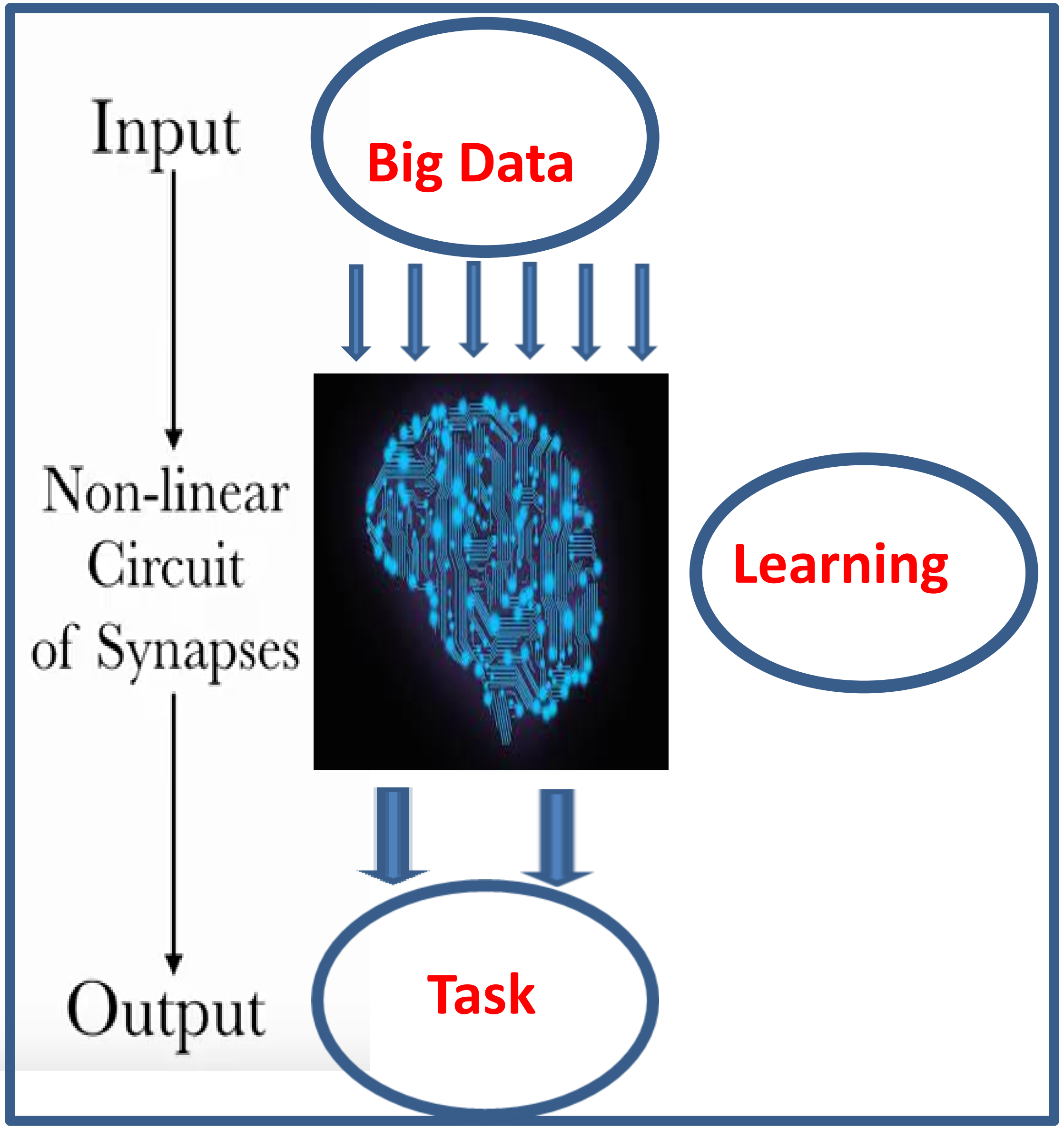
Microscopic systems made of many interacting particles in **chemistry**, **materials science**, **atomic and molecular physics**

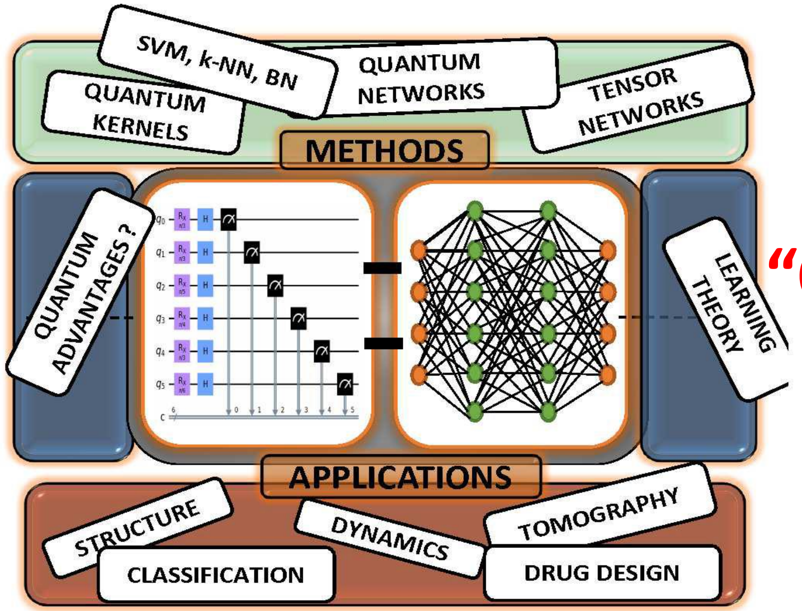
Quantum Mechanics has to be used to provide an accurate description of the system





Quantum Machine-Learning





Review Article

"Quantum Machine Learning for Chemistry and Physics"

Manas Sajjan, Junxu Li, Raja Selvarajan, Shree Hari Sureshbabu ,Sumit Suresh Kale, Rishabh Gupta, Vinit Singh and Sabre Kais*

Department of Chemistry, Purdue University

Chemical Society Reviews 51, 6475 – 6573 (2022)

- **QML Model:** **Restricted Boltzmann Machine (RBM)**
- **Implementation on a Quantum Device:** Designing a quantum circuit with quadratic resource requirements (circuit width, circuit depth, parameter count)
  - **Applications:** Electronic structure of simple molecules and 2-D materials such as Hexagonal Boron Nitride, Graphene, Molybdenum Disulfide MoS2, Tungsten disulfide WS_2, Emerging Phenomena in quantum and topological materials.

Focus on : Statistical Physics and Machine Learning

- **Information Scrambling:** Perspective and Out-of-Time-Order Correlator (OTOC) and the learning landscape of a quantum machine learning algorithm.

Electronic Structure of Molecules and Materials on Quantum Computers

Challenge

Describing the nontrivial correlations encoded in the exponential complexity of the many-body wave function

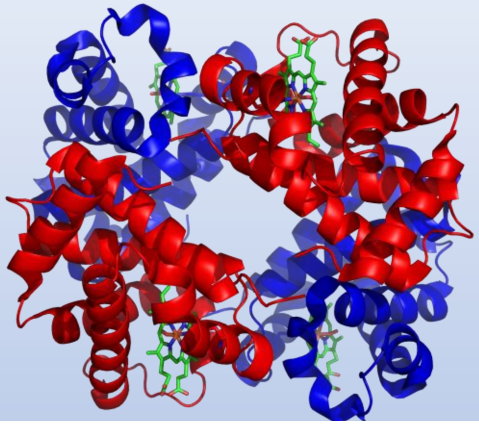
Goal: Understanding the properties of quantum many-body systems

Hilbert Space

Exponential Complexity

N-Spins 2^N configurations

N=80, Avogadro 10^{23}



Average size of human protein is 300 amino acid residues

20 base amino acids, so there are 20^{300} possible sequences

10^{390}

the number of different existing proteins in all living organisms

10^{7}

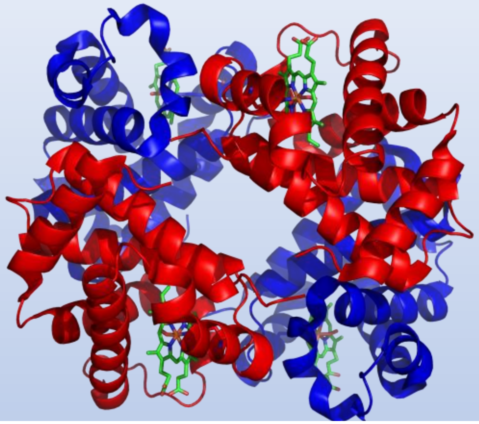
Goal: **Understanding the properties of quantum many-body systems**

Hilbert Space

Exponential Complexity

N-Spins 2^N configurations

N=80, Avogadro 10^{23}



Average size of human protein is 300 amino acid residues

20 base amino acids, so there are 20^{300} possible sequences

10^{390}

the number of different existing proteins in all living organisms

10^{7}

Physical states usually only access a tiny corner of the entire Hilbert space

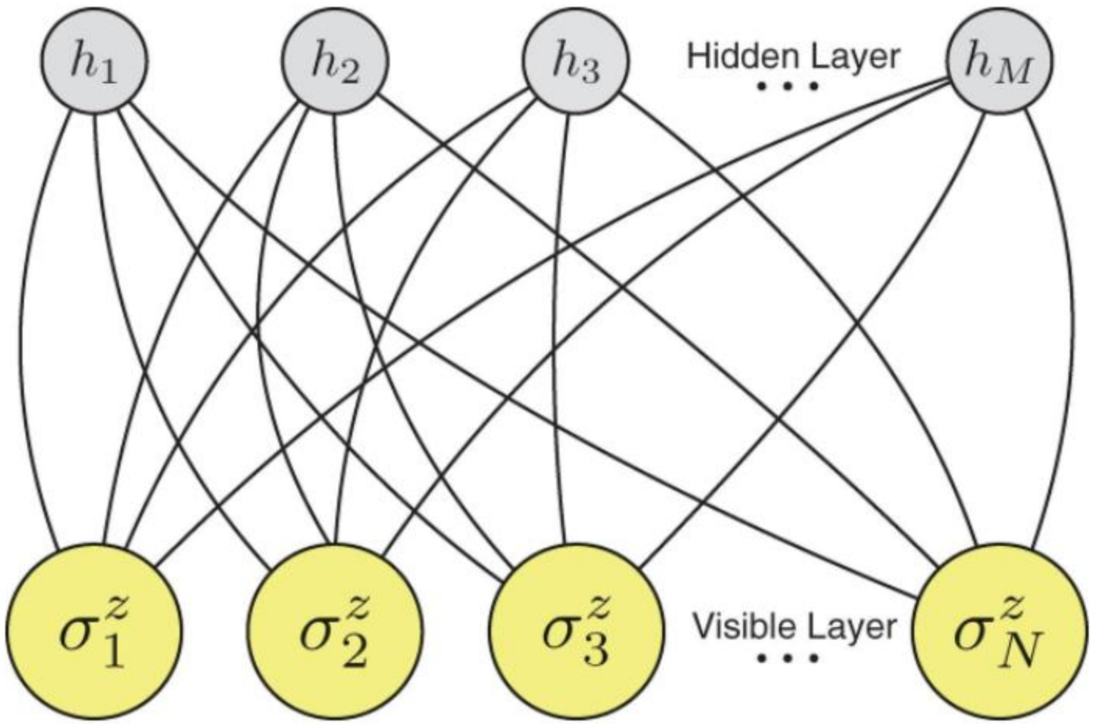
Quantum Machine-Learning

Deep Learning and Artificial Neural Network
"Restricted Boltzmann Machine (RBM)"

c_i →

W_{ij} →

b_j →



$$E_\lambda(\mathbf{v},\mathbf{h}) = -\sum_{i,j} W_{ij} h_i v_j - \sum_{j=1}^{V} b_j v_j - \sum_{i=1}^{H} c_i h_i$$

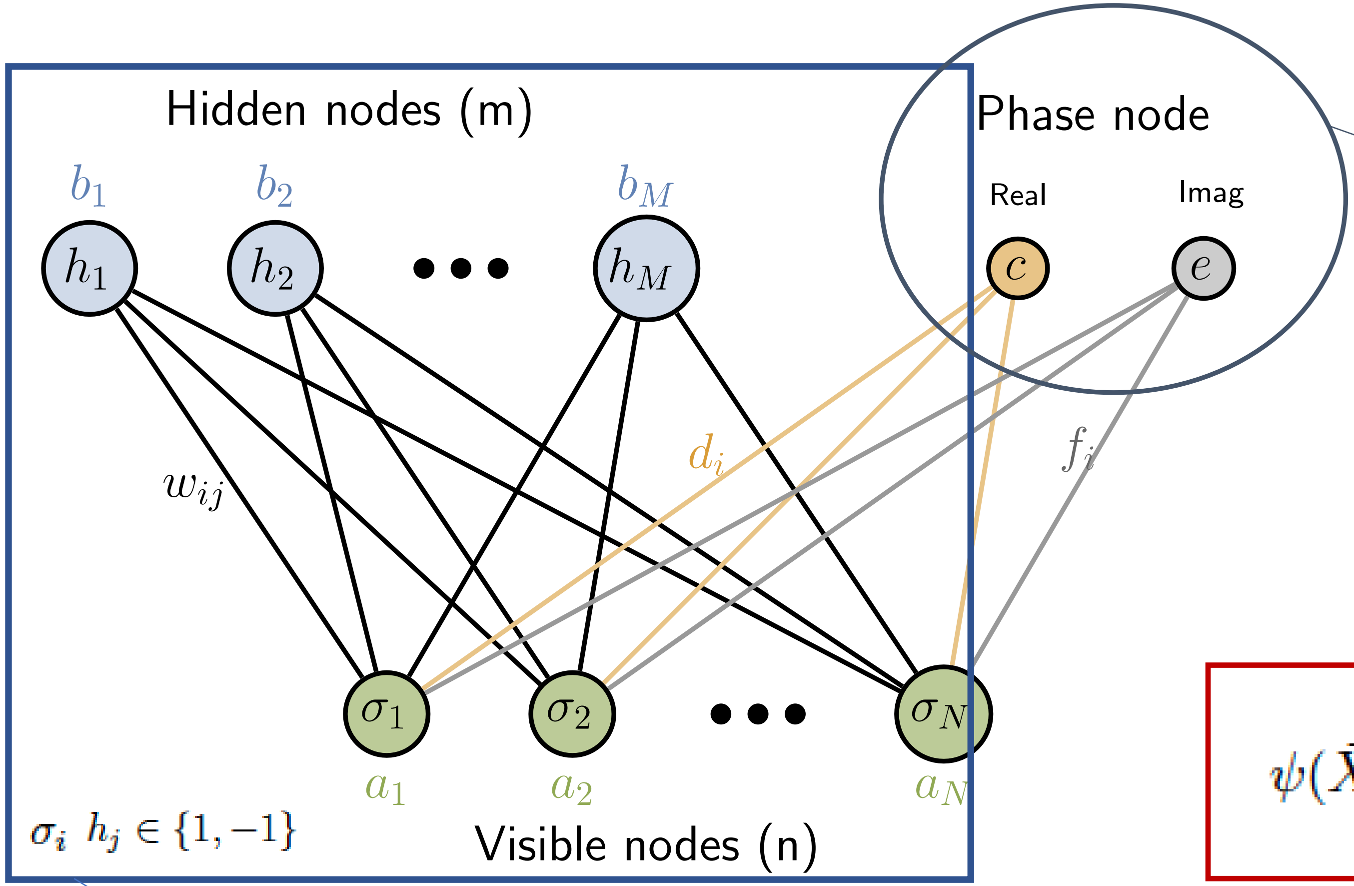
$$p_\lambda(\mathbf{v},\mathbf{h}) = \frac{1}{Z_\lambda} e^{-E_\lambda}$$

$$p(\mathbf{v}) = \sum_{\mathbf{h}} p(\mathbf{v},\mathbf{h})$$

wavefunction $p(\mathbf{v}) = |\langle \mathbf{v}|\Psi\rangle|^2$

R.G. Melco et al. Nature Physics, 15, 887 (2019)
Carleo & Troyer, Science 355, 602 (2017)

NETWORK ARCHITECTURE (RESTRICTED BOLTZMANN MACHINE ANSATZ)



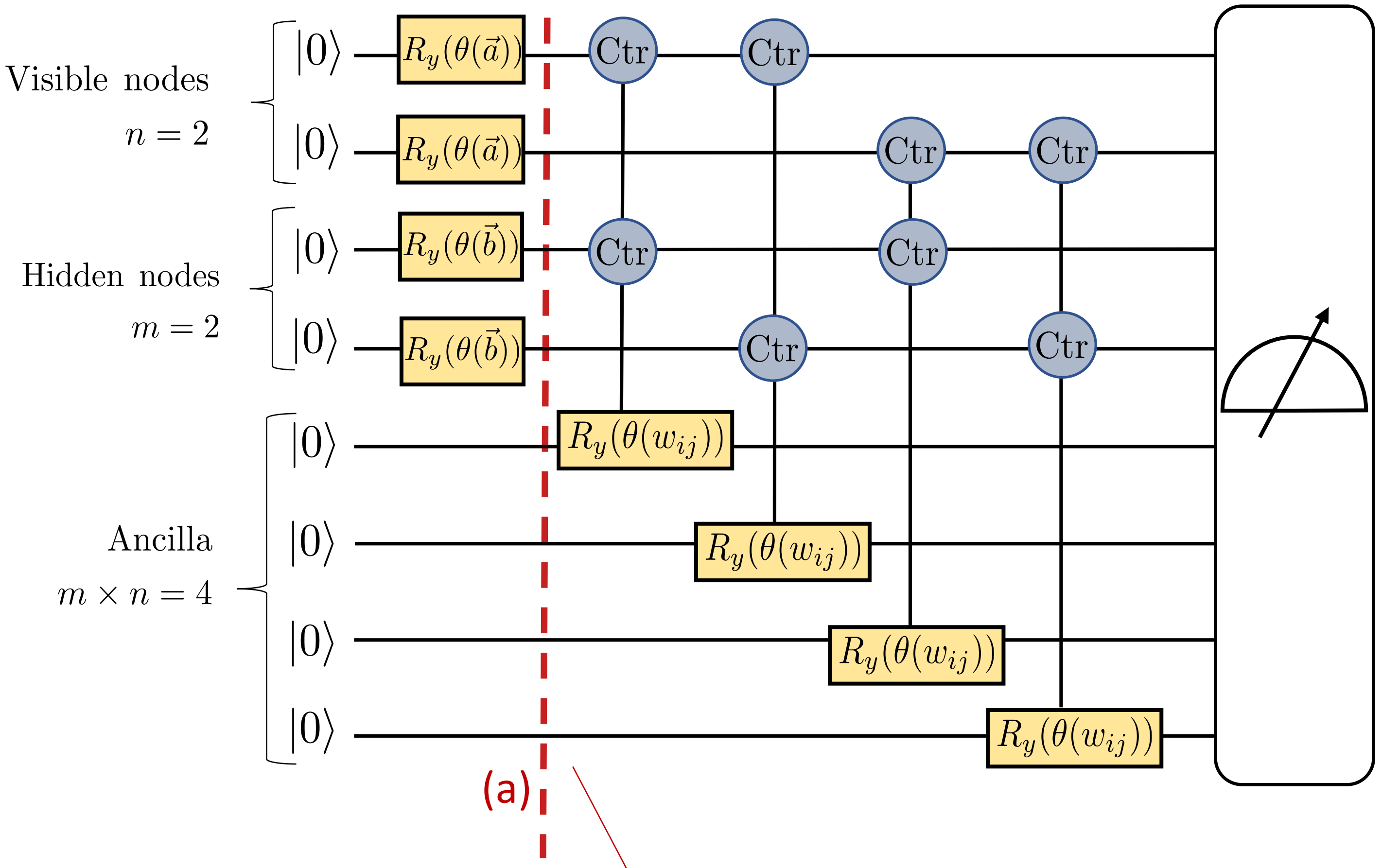
$$s(\vec{d}, \vec{f}, c, e, \vec{\sigma}) = \tanh\left[(c + \sum_i d_i \sigma_i) + i(e + \sum_i f_i \sigma_i)\right]$$

$$\psi(\vec{X}) = \sum_{\sigma} \sqrt{\tilde{p}(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma})} s(\vec{d}, \vec{f}, c, e, \vec{\sigma}) |\sigma_1 \sigma_2 \ldots \sigma_n\rangle$$

$$P(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma}, \vec{h}) = \frac{e^{\sum_i a_i \sigma_i + \sum_j b_j h_j + \sum_{ij} w_{ij} \sigma_i h_j}}{\sum_{\{\sigma h\}} e^{\sum_i a_i \sigma_i + \sum_j b_j h_j + \sum_{ij} w_{ij} \sigma_i h_j}}$$

$$\tilde{p}(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma}) = \sum_h P(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma}, \vec{h})$$

CONSTRUCTION OF AMPLITUDE USING QUANTUM CIRCUIT



$$R_y(\theta(\vec{a})) = \begin{bmatrix} \sqrt{\frac{e^{-a_i/k}}{e^{-a_i/k}+e^{a_i/k}}} & \sqrt{\frac{e^{a_i/k}}{e^{-a_i/k}+e^{a_i/k}}} \\ \sqrt{\frac{e^{a_i/k}}{e^{-a_i/k}+e^{a_i/k}}} & \sqrt{\frac{e^{-a_i/k}}{e^{-a_i/k}+e^{a_i/k}}} \end{bmatrix}$$

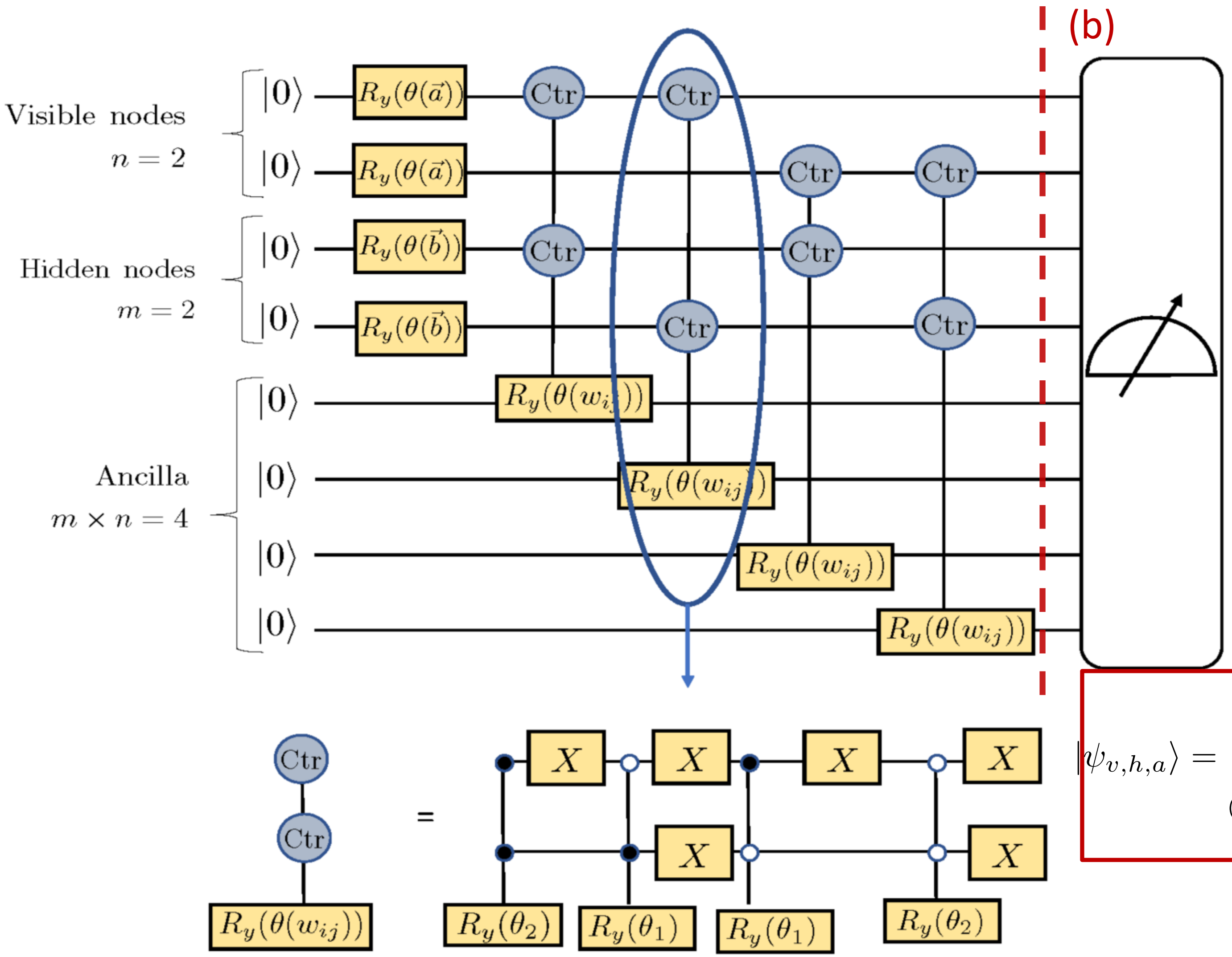
$$R_y(\theta(\vec{b})) = \begin{bmatrix} \sqrt{\frac{e^{-b_j/k}}{e^{-b_j/k}+e^{b_j/k}}} & \sqrt{\frac{e^{b_j/k}}{e^{-b_j/k}+e^{b_j/k}}} \\ \sqrt{\frac{e^{b_j/k}}{e^{-b_j/k}+e^{b_j/k}}} & \sqrt{\frac{e^{-b_j/k}}{e^{-b_j/k}+e^{b_j/k}}} \end{bmatrix}$$

State at (a)

$$|\psi_{v,h,a}\rangle = \sum_{(\vec{\sigma},\vec{h})} \sqrt{O(\vec{\sigma},\vec{h},\vec{a},\vec{b})}|\vec{\sigma}\vec{h}\rangle_{vh} \otimes |0\rangle_a$$

$$O(\vec{\sigma},\vec{h},\vec{a},\vec{b}) = \frac{e^{\frac{\sum_i a_i\sigma_i+\sum_j b_j h_j}{k}}}{\sum_{\vec{\sigma}\vec{h}} e^{\frac{\sum_i a_i\sigma_i+\sum_j b_j h_j}{k}}}$$

CONSTRUCTION OF AMPLITUDE USING QUANTUM CIRCUIT



$$CCR_y(\theta_2(\vec{W}))|11\rangle_{vh}|0\rangle_a = \sqrt{\frac{e^{W_{ij}/k}}{e^{|W_{ij}|/k}}}|11\rangle_{vh}|1\rangle_a + \sqrt{1-\frac{e^{W_{ij}/k}}{e^{|W_{ij}|/k}}}|11\rangle_{vh}|0\rangle_a$$

$$CCR_y(\theta_2(\vec{W}))|00\rangle_{vh}|0\rangle_a = \sqrt{\frac{e^{W_{ij}/k}}{e^{|W_{ij}|/k}}}|00\rangle_{vh}|1\rangle_a + \sqrt{1-\frac{e^{W_{ij}/k}}{e^{|W_{ij}|/k}}}|00\rangle_{vh}|0\rangle_a$$

$$CCR_y(\theta_1(\vec{W}))|10\rangle_{vh}|0\rangle_a = \sqrt{\frac{e^{-W_{ij}/k}}{e^{|W_{ij}|/k}}}|10\rangle_{vh}|1\rangle_a + \sqrt{1-\frac{e^{-W_{ij}/k}}{e^{|W_{ij}|/k}}}|10\rangle_{vh}|0\rangle_a$$

$$CCR_y(\theta_1(\vec{W}))|01\rangle_{vh}|0\rangle_a = \sqrt{\frac{e^{-W_{ij}/k}}{e^{|W_{ij}|/k}}}|01\rangle_{vh}|1\rangle_a + \sqrt{1-\frac{e^{-W_{ij}/k}}{e^{|W_{ij}|/k}}}|01\rangle_{vh}|0\rangle_a$$

State at (b)

$$|\psi_{v,h,a}\rangle = \sum_{(\vec{\sigma},\vec{h})} \sqrt{O(\vec{\sigma},\vec{h},\vec{a},\vec{b})}|\vec{\sigma}\vec{h}\rangle_{vh} \otimes \sqrt{(1-\eta(\vec{W},\vec{\sigma},\vec{h})}|0\rangle_a + \sqrt{\eta(\vec{W},\vec{\sigma},\vec{h})}|1\rangle_a)$$

$$\eta(\vec{W},\vec{\sigma},\vec{h}) = \frac{e^{\frac{1}{k}(\sum_{i,j} w_{ij}\sigma_i h_j)}}{e^{\frac{1}{k}\sum_{i,j}|w_{ij}|}}$$

Methodology: Summary

- **The wavefunction can be expressed as:**

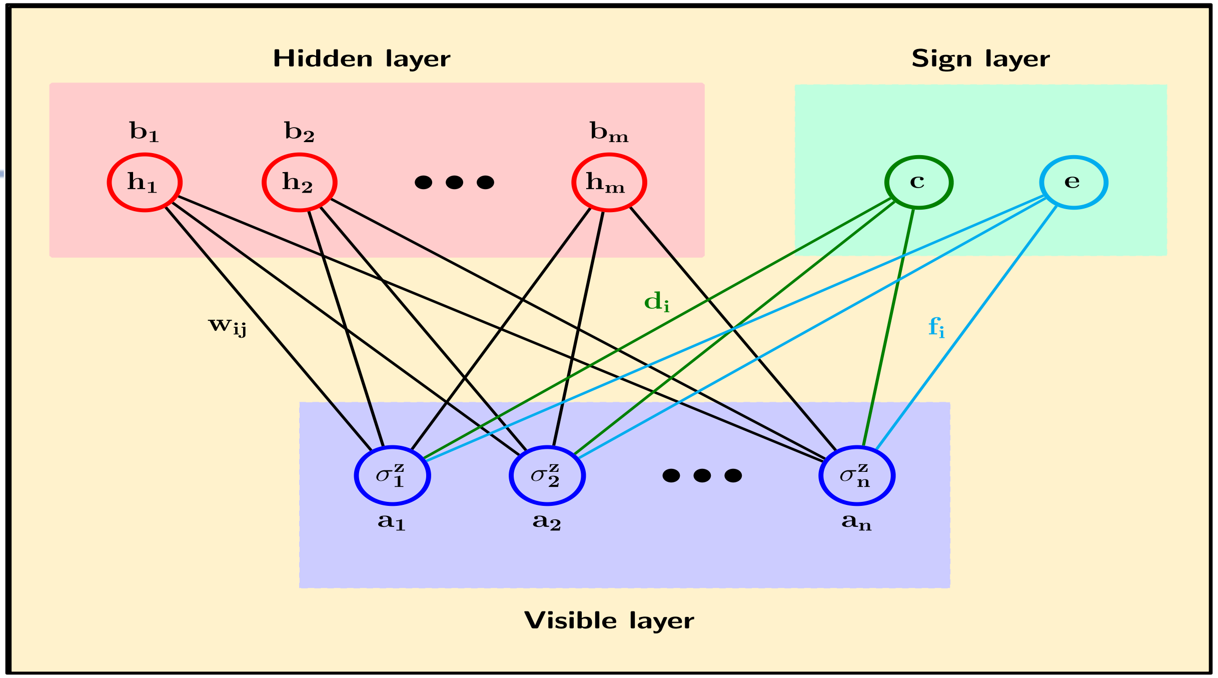
$$|\phi\rangle = \sum_x \sqrt{P(x)}\, s(x)|x\rangle$$

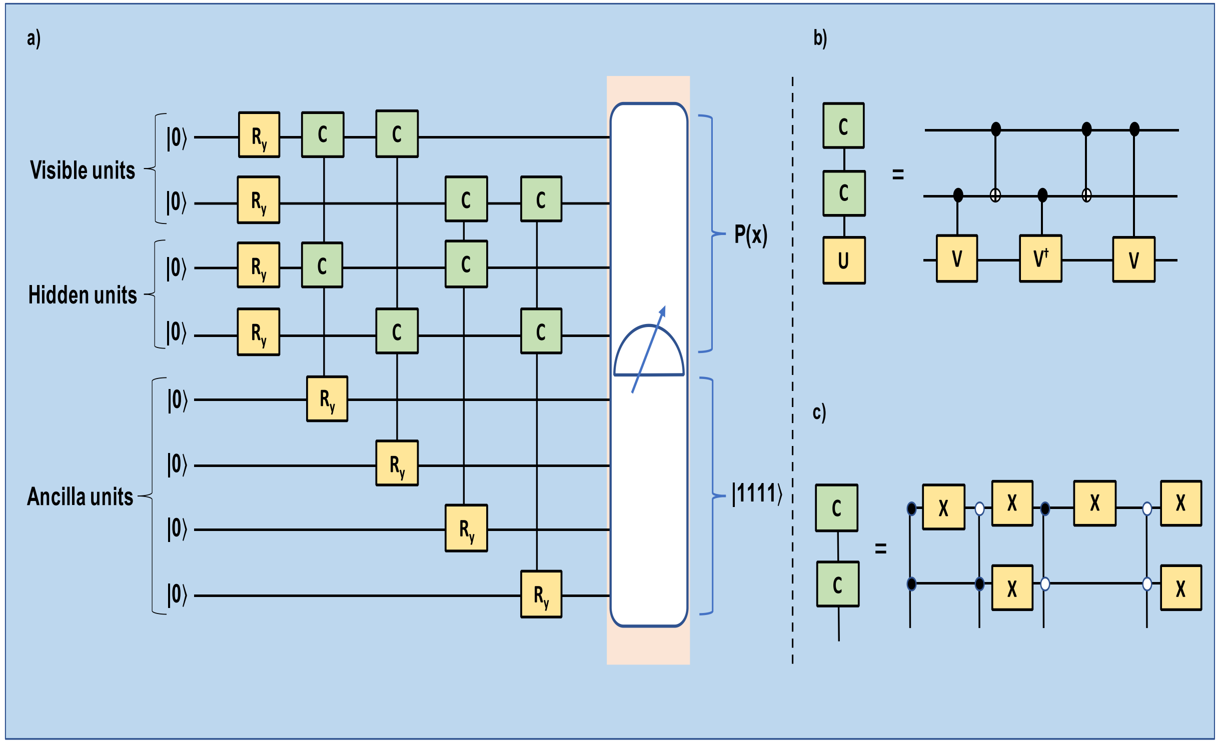
$$P(\mathbf{x}) = \frac{\sum_{\{h\}} e^{\sum_i a_i\sigma_i^z + \sum_j b_j h_j + \sum_{ij} w_{ij}\sigma_i^z h_j}}{\sum_{x'} \sum_{\{h\}} e^{\sum_i a_i\sigma_i^z + \sum_j b_j h_j + \sum_{ij} w_{ij}\sigma_i^z h_j}}$$

$$s(x) = tanh\left(\left(c + \sum_i d_i\, \sigma_i^z\right) + i(e + \sum_i f_i\, \sigma_i^z)\right)$$

- **Given Hamiltonian H and a trial state $|\phi\rangle = \sum_x \phi(\mathbf{x})\, |x\rangle$** we compute the expectation value:

$$\langle H\rangle = \frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle}$$





Resource Requirements

No. of qubits	**No. of gates**	**Variational parameter count**
$(m + n + mn) = O(mn)$	$(m + n + mn) = O(mn)$	$(m + 3n + mn + 2) = O(mn)$

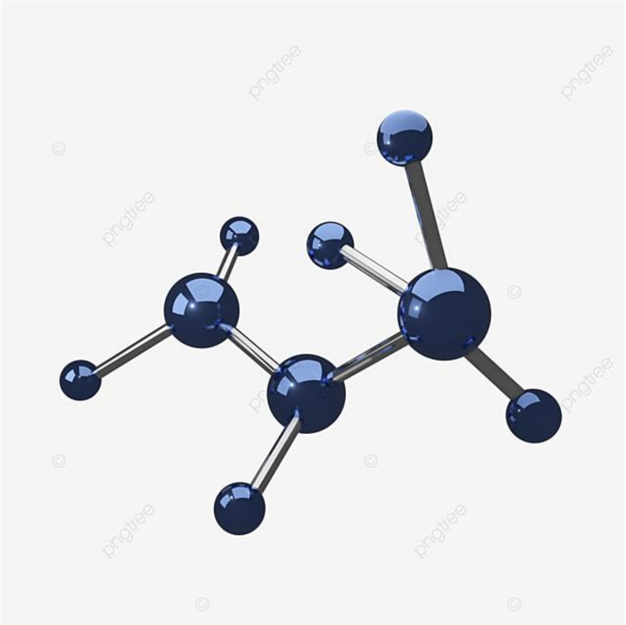
Simulating the full RBM distribution classically will always likely be **Exponential** but on a quantum computer will be **Quadratic Resources.**

Also RBM is a universal approximator for any probability density

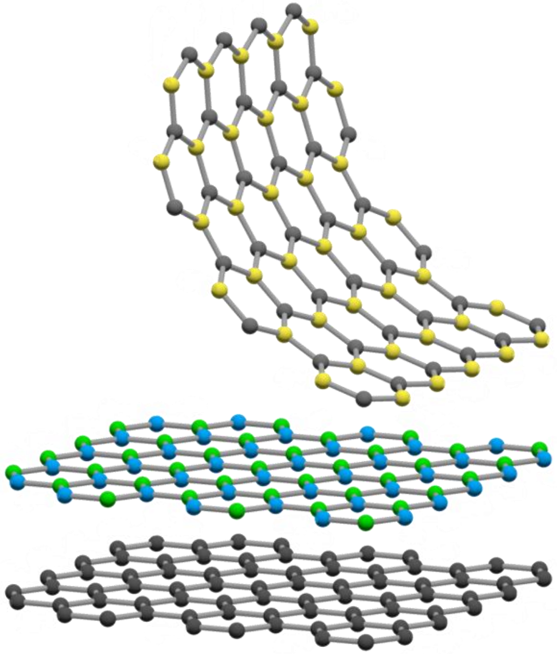
Long, P. M.; Servedio, R. A. Restricted Boltzmann Machines are hard to approximately evaluate or simulate. ICML 2010 - Proceedings, 27th International Conference on Machine Learning **2010**, 703–710.

SYSTEM OF INTEREST (DRIVER)

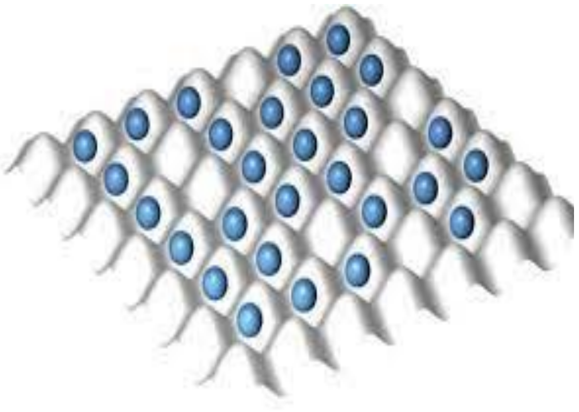
Could be a material, a molecule or any spin-lattice models



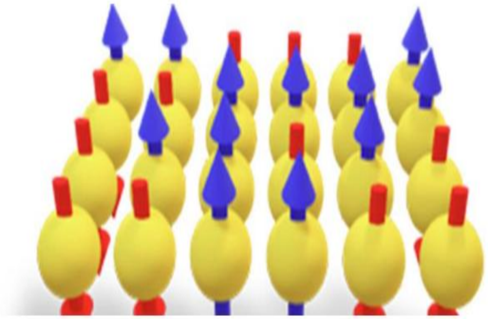
Molecules



Materials

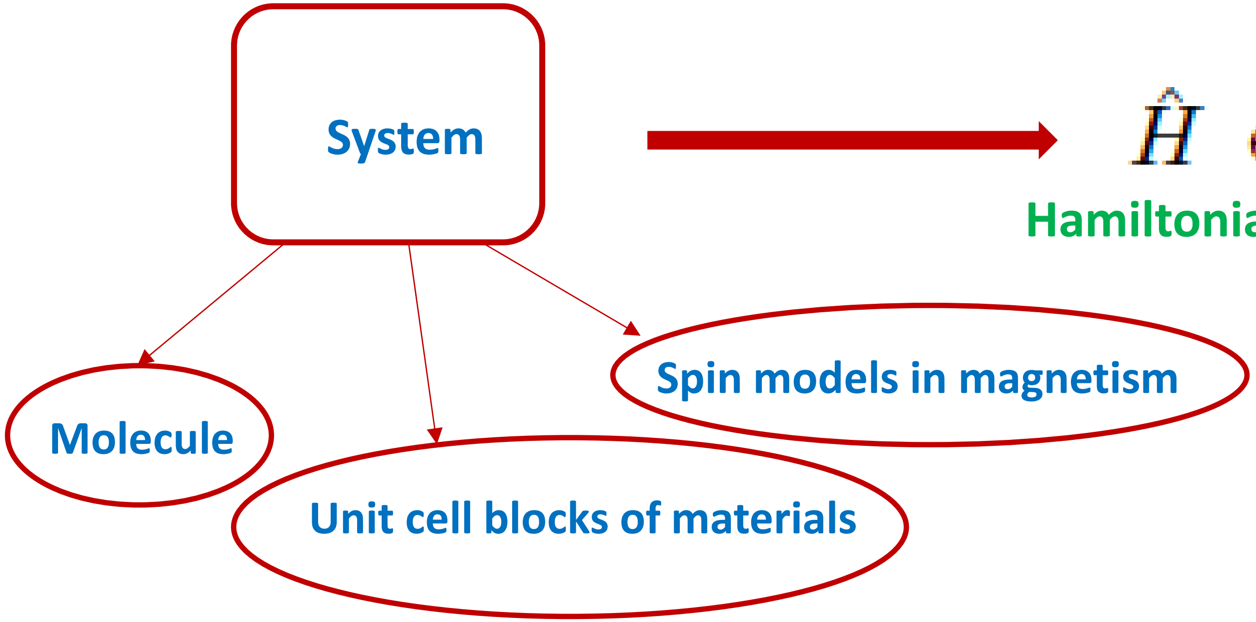


Ultracold Atomic-Lattice



Spin-Lattice Models

OBJECTIVE



$$\hat{H} \in \mathbb{C}^{d\times d}$$

Hamiltonian matrix for the system in a chosen representation

DEFINITION OF THE PROBLEM

- Construct approximate eigenstates $\psi(\vec{X})$ of $\hat{H}$ without the prohibitively expensive exact diagonalization

- Filter eigenstates with symmetry operators based on user defined choices. This means if the model admits then

$$\{\hat{O}_i\}_{i=1}^{i=k}, \ s.t \ [\hat{H}, \hat{O}_i] = 0 \ \forall i \in \{1, 2, ...k\}$$

grouping eigenstates based on eigenspaces of operators specified by the user

BACKBONE OF THE TALK

nature COMMUNICATIONS

***Nature Comm.* 9, 4195 (2018)**

ARTICLE

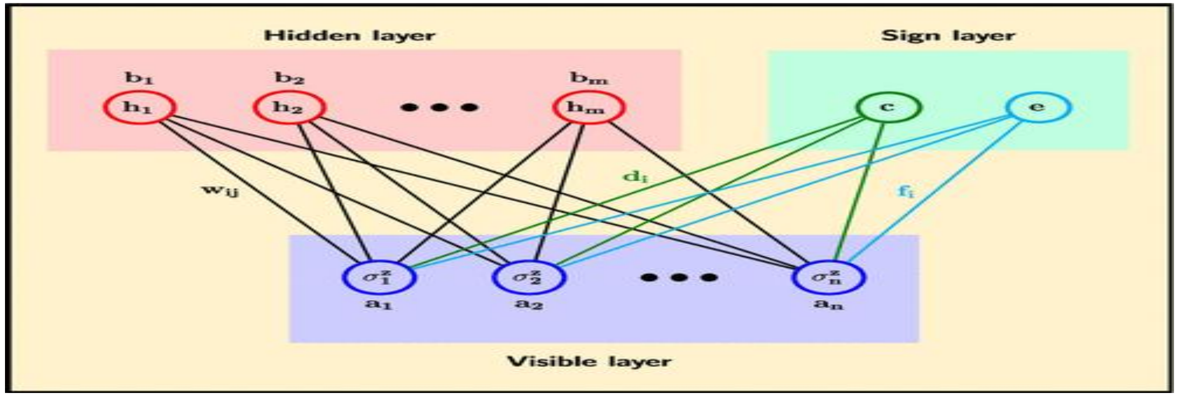
DOI: 10.1038/s41467-018-06598-z OPEN

Quantum machine learning for electronic structure calculations

Rongxin Xia[1] & Sabre Kais[1,2,3]

Implementation of Quantum Machine Learning for Electronic Structure Calculations of Periodic Systems on Quantum Computing Devices

S.H Sureshbabu, M. Sajjan, S. Oh and S. Kais



Journal of Chemical Information and Modeling, 61, 2667 (2021)

J|A|C|S
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

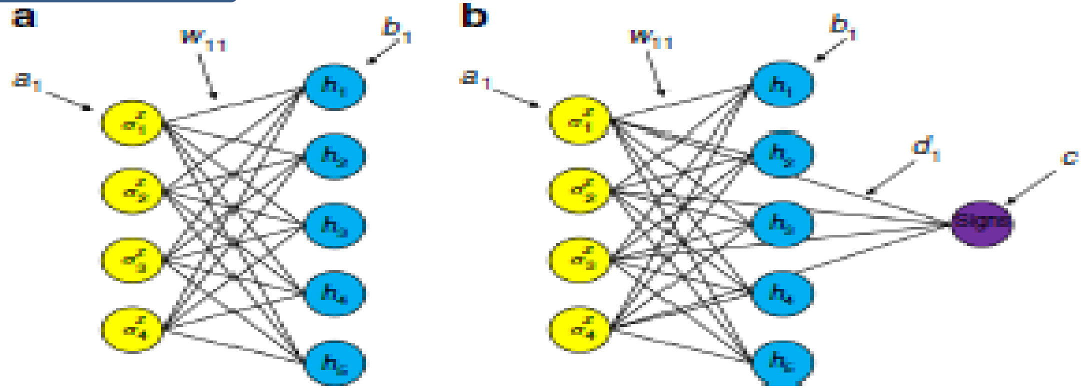
pubs.acs.org/JACS

***143* (44), 18426 October (2021)**

Quantum Machine-Learning for Eigenstate Filtration in Two-Dimensional Materials

Manas Sajjan, Shree Hari Sureshbabu, and Sabre Kais*

Results



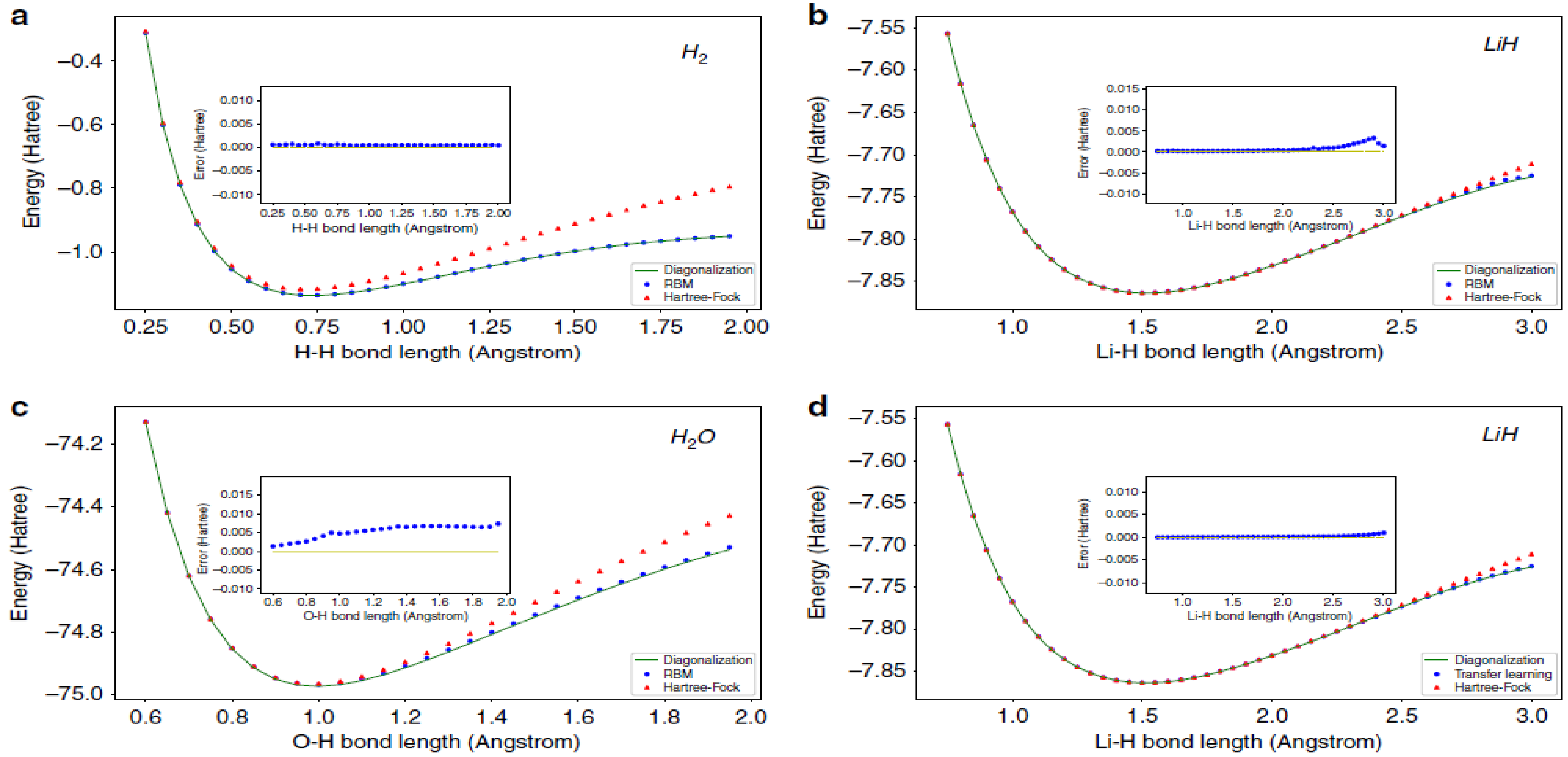
***Nature Comm.* 9, 4195 (2018)**

RTICLE

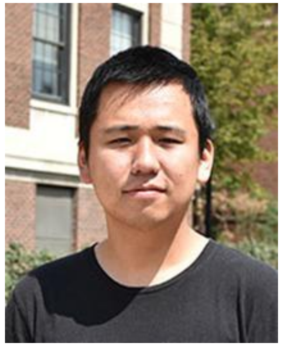
I: 10.1038/s41467-018-06598-z OPEN

Quantum machine learning for electronic tructure calculations

ngxin Xia[1] & Sabre Kais[1,2,3]



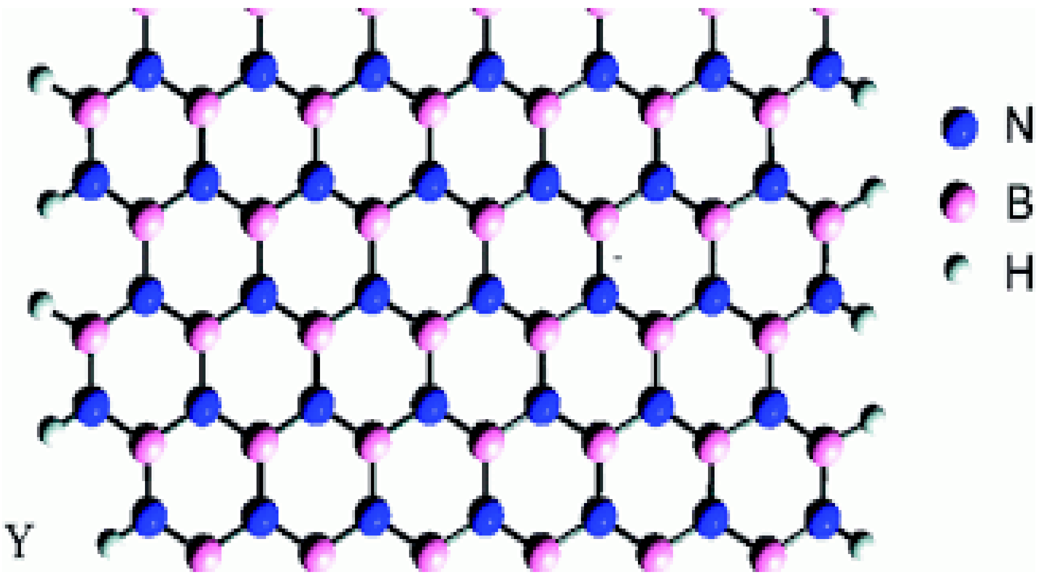
Rongxin Xia
Facebook



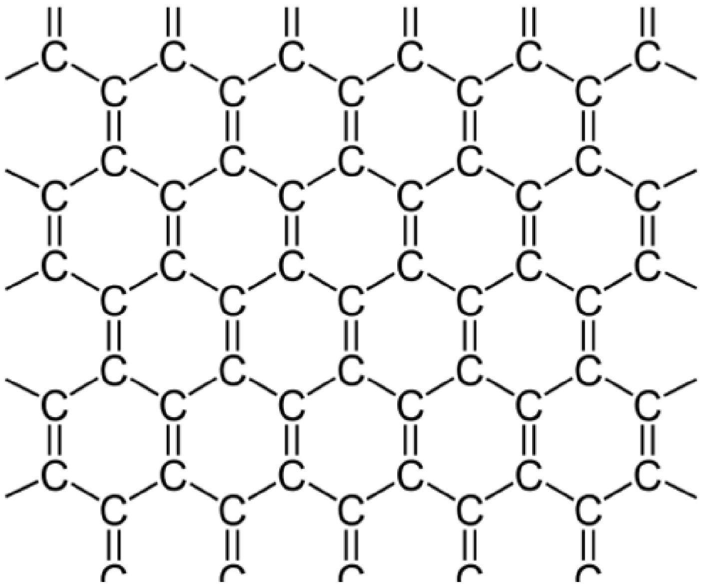
The results of

H_2 (n = 4, m =8, iterations=10,000), LiH (n = 4, m = 8, iterations=40,000) and H_2O (n = 6, m = 6, iterations=40,000)

Hexagonal-Boron Nitride (h-BN)



Graphene



Shree Hari

JP Morgan Chase & Co.

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Journal of Chemical Information and Modeling, 61. 2667 (2021)

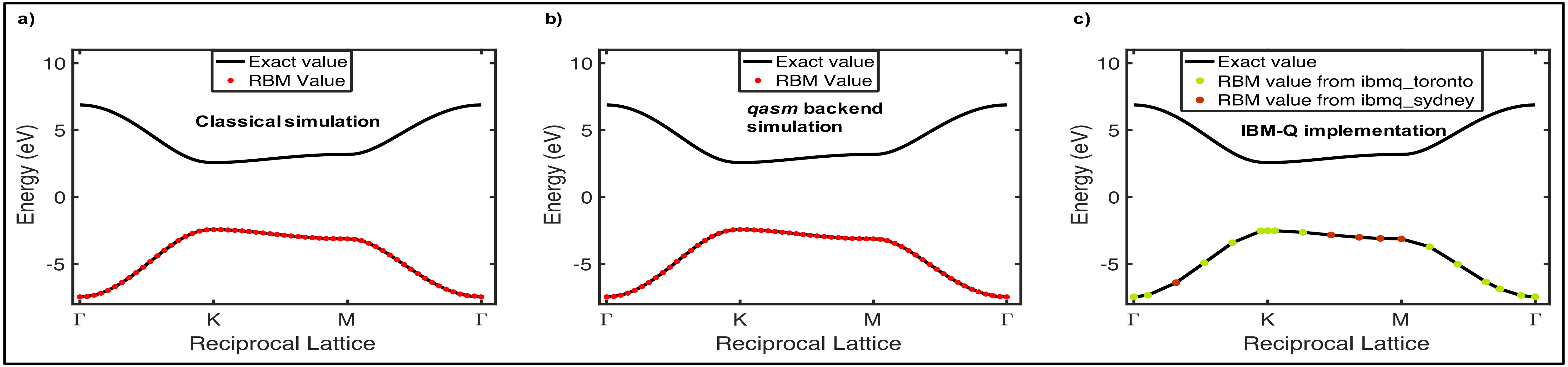
pubs.acs.org/jcim Article

Implementation of Quantum Machine Learning for Electronic Structure Calculations of Periodic Systems on Quantum Computing Devices

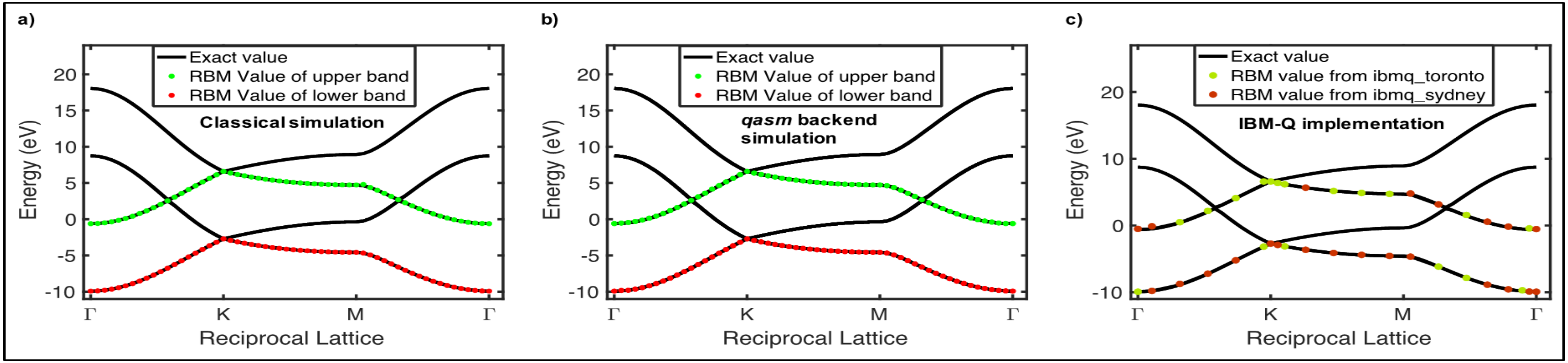
Shree Hari Sureshbabu, Manas Sajjan, Sangchul Oh, and Sabre Kais*

Results

Implementation on IBM-Q hexagonal Boron Nitride



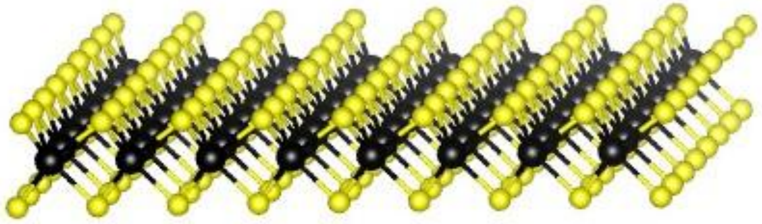
Graphene



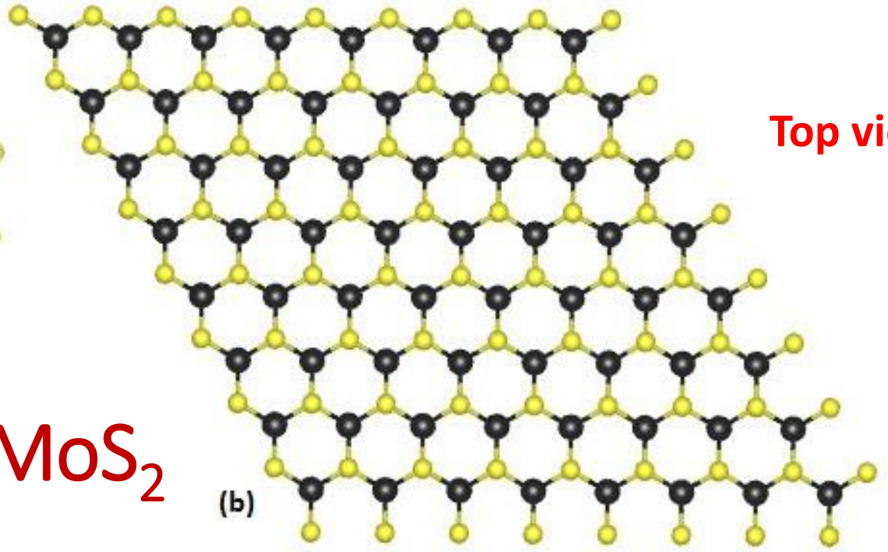
Band Structure of 2-D materials

(a)

Side view



Molybdenum disulfide MoS_2



(b)

Top view

Tungsten disulfide WS_2

Dr. Manas Sajjan

J | A | C | S
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

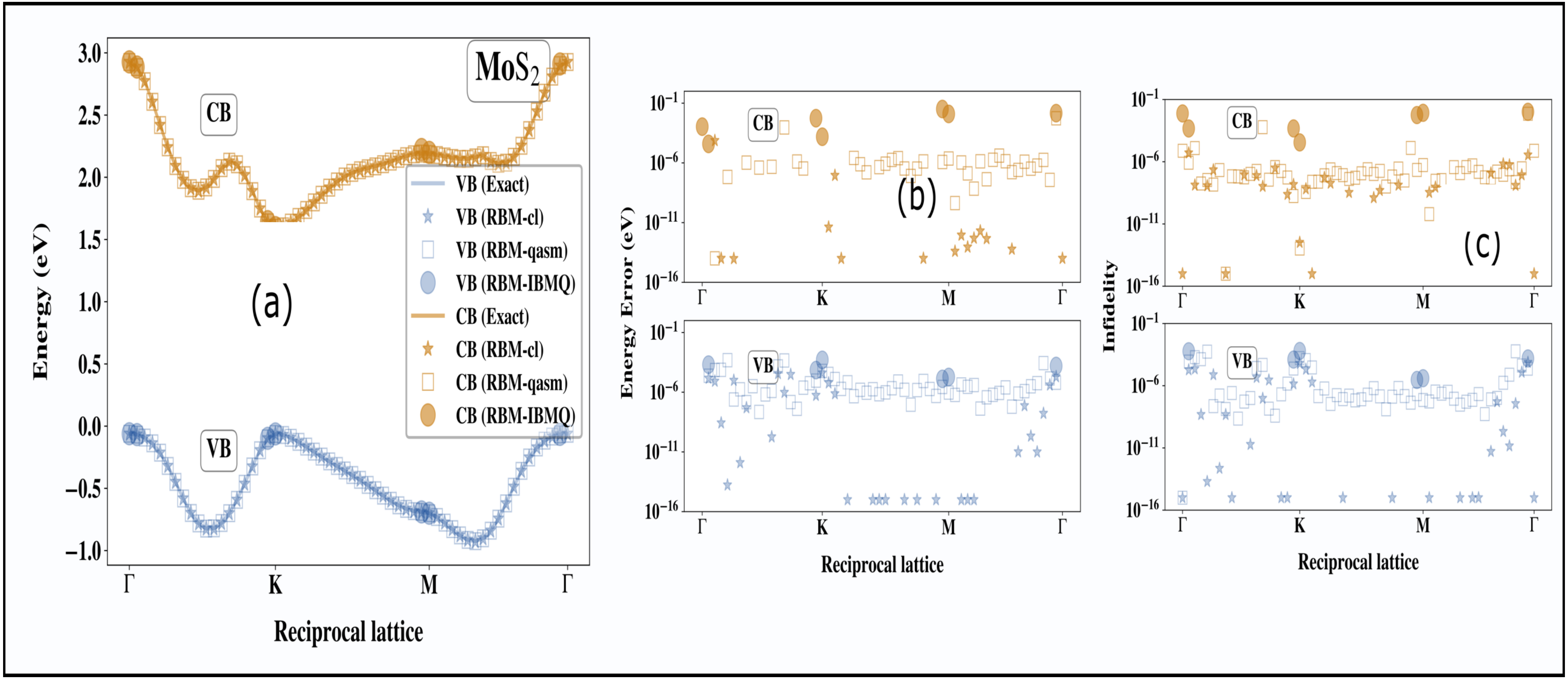
pubs.acs.org/JACS Article

Quantum Machine-Learning for Eigenstate Filtration in Two-Dimensional Materials

Manas Sajjan, Shree Hari Sureshbabu, and Sabre Kais*

Cite This: https://doi.org/10.1021/jacs.1c06246 Read Online

VALENCE AND CONDUCTION BAND - RESULTS FOR MoS_2



Deeper Understanding of the Learning Process

(Connection to Statistical Physics)

Out-of-time-order correlator (OTOC) and scrambling of quantum information

Navigating the learning landscape of a Quantum Machine Learning model through rigorous correlation bounds

Dr. Manas Sajjan

Information Scrambling Perspective
and
Out-of-Time-Order Correlator (OTOC)

Vinit Singh

PHYSICAL REVIEW RESEARCH 5, 013146 (2023)

- How neural network exchanges information among sub-units and **how fast an initial excitation travel** using statistical correlators ?

- What is the **connection** of the above point to **information theory** ? Is there an information bottleneck which can affect trainability ?

- What are the **universal features** of such exchange of information ?
- Can we expedite training and **construct better networks** using this

Imaginary components of out-of-time-order correlator and information scrambling for navigating the learning landscape of a quantum machine learning model

Manas Sajjan,[1,4,*] Vinit Singh,[1,4,*] Raja Selvarajan,[2,4] and Sabre Kais[1,2,3,4,†]
[1]Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA
[2]Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, USA
[3]Department of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA
[4]Purdue Quantum Science and Engineering Institute, Purdue University, West Lafayette, Indiana 47907, USA

(Received 30 August 2022; revised 13 January 2023; accepted 23 January 2023; published 27 February 2023)

We introduce and analytically illustrate that hitherto unexplored imaginary components of out-of-time order correlators can provide unprecedented insight into the information scrambling capacity of a graph neural network. Furthermore, we demonstrate that it can be related to conventional measures of correlation like quantum mutual information and rigorously establish the inherent mathematical bounds (both upper and lower bound) jointly shared by such seemingly disparate quantities. To consolidate the geometrical ramifications of such bounds during the dynamical evolution of training we thereafter construct an emergent convex space. This newly designed space offers much surprising information including the saturation of lower bound by the trained network even for physical systems of large sizes, transference, and quantitative mirroring of spin correlation from the simulated physical system across phase boundaries as desirable features within the latent subunits of the network (even though the latent units are directly oblivious to the simulated physical system) and the ability of the network to distinguish exotic spin connectivity (volume law vs area law). Such an analysis demystifies the training of quantum machine learning models by unraveling how quantum information is scrambled through such a network introducing correlation surreptitiously among its constituent subsystems and open a window into the underlying physical mechanism behind the emulative ability of the model.

OUT-OF-TIME-ORDER CORRELATORS (OTOC)

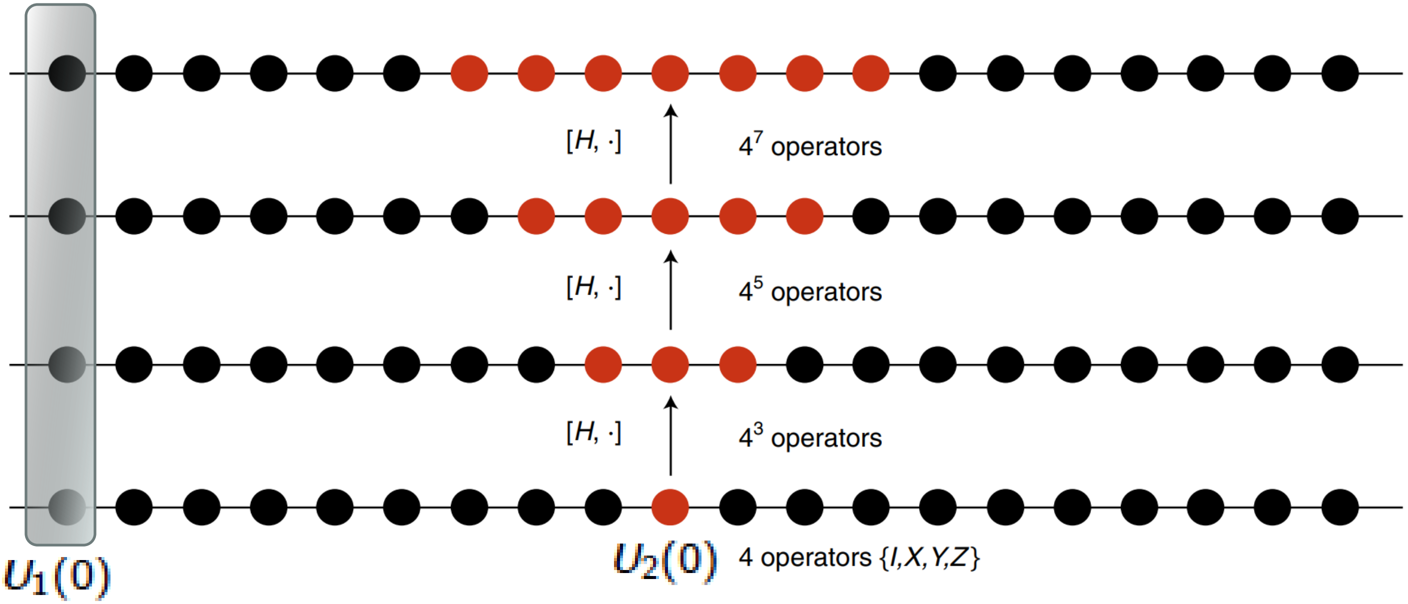
A four-point temporal correlation function defined between sites (1,2) with local unitaries

$U_1(0)\ \ U_2(0)$

$$C_{U_1,U_2}(t) = \langle U_1^\dagger(0) U_2^\dagger(t) U_1(0) U_2(t) \rangle$$

$$U_2(t) = e^{iHt} U_2(0) e^{iHt}$$

$$= \sum_{k=0}^{\infty} \frac{(it)^k}{k!} [H, [H, \cdots [H, U_2(0)] \cdots]]$$



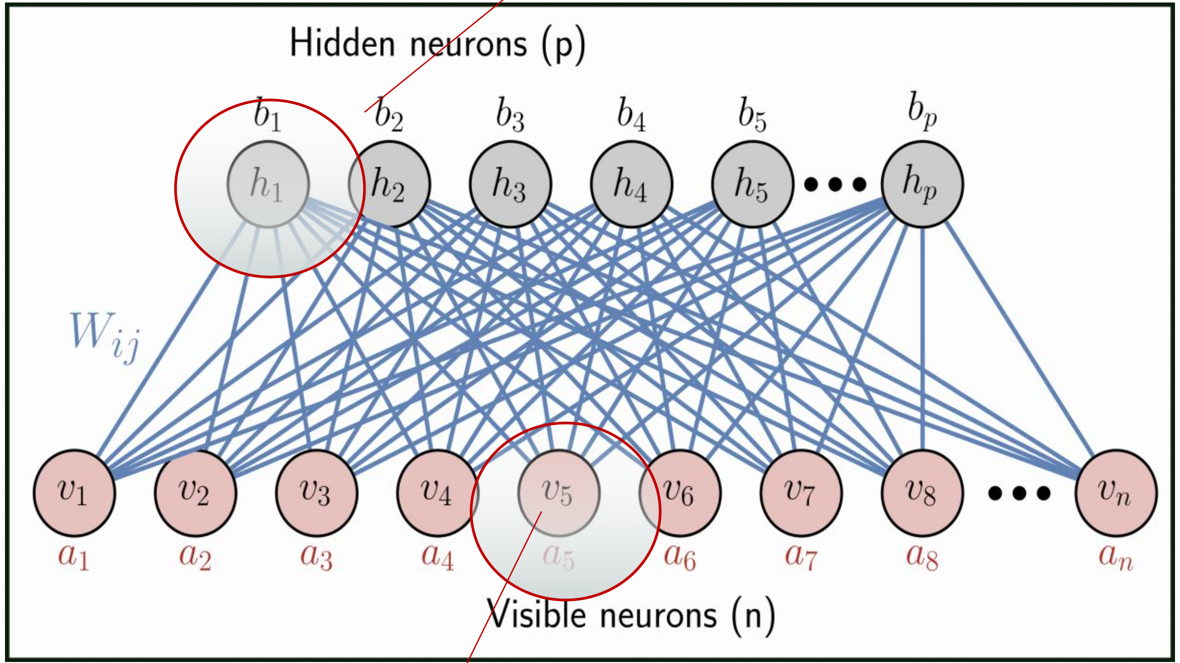
- If the sites (1,2) are far apart then the supporting bases of operators $U_1(0)$ and $U_2(0)$ then $C_{U_1,U_2}(0)$ will be 1
- However, with time as the supporting bases of $U_2(t)$ grows then the overlap between the two operators will increase leading to different $C_{U_1,U_2}(t)$
- $C_{U_1,U_2}(t)$ is a measure of how fast the information has travelled leading to the overlap under the effect of interactions in the system

Brian Swingle

Nature Physics | VOL 14 | OCTOBER 2018 | 988–990 |

OTOC FOR THE NETWORK

$$U_2(0) = \vec{\sigma}_\beta(h_m, 0) - \kappa_2 \mathcal{I}$$



$U_2(0)$ evolves in time under the effect of

$$\mathcal{H}(\vec{X}, \vec{v}, \vec{h}) = \sum_{i=1}^{n} a_i \sigma^z(v_i) + \sum_{j=1}^{p} b_j \sigma^z(h_j) + \sum_{i=1, j=1}^{n,p} W_j^i \sigma^z(v_i) \sigma^z(h_j)$$

$$U_1(0) = \sigma_\alpha(v_k, 0) - \kappa_1 \mathcal{I}$$

$$C_{\sigma^\alpha, \sigma^\beta}(\kappa_1, \kappa_2, \vec{X}, t) = \langle \tilde{\sigma}^\alpha(v_k, 0) \tilde{\sigma}^\beta(h_m, t) \tilde{\sigma}^\alpha(v_k, 0) \tilde{\sigma}^\beta(h_m, t) \rangle$$

$$C_{km}^{km}(\vec{X},t) = Cos(4W_m^k t) \; + iCov(\sigma_k^x(v), \sigma_m^x(h))Sin(4W_m^k t)$$

$$Cov(\sigma_k^x(v), \sigma_m^x(h)) = \langle \sigma_k^x(v), \sigma_m^x(h) \rangle_{\rho^{th}} \; - \langle \sigma_k^x(v) \rangle_{\rho^{th}} \langle \sigma_k^x(h) \rangle_{\rho^{th}}$$

Mutual Information

$$\mathcal{I}(v_k, h_m) = S(^1\rho(v_k)) + S(^1\rho(h_m)) - S(^2\rho(v_k, h_m))$$

Von-Neumann Entropies

$$S(Y) = -Tr(Y ln Y).$$

$$LB \leq \mathcal{I}(v_k, h_m) \leq UB$$

wherein the lower bound (LB) is

$$LB = 2 + \frac{1+\eta(\vec{X})}{2} + ln(\frac{1+\eta(\vec{X})}{4})$$
$$+ \frac{1-\eta(\vec{X})}{2} + ln(\frac{1-\eta(\vec{X})}{4})$$

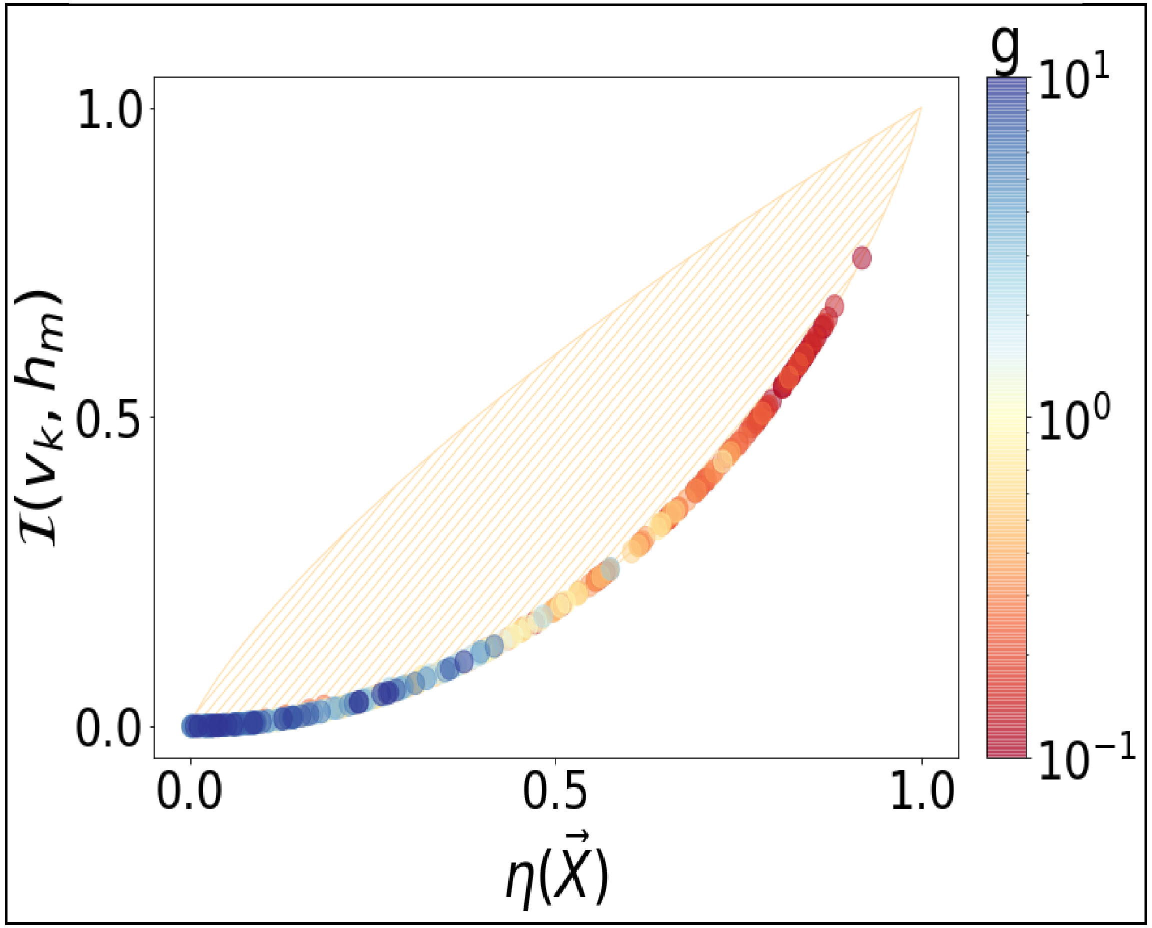
and the corresponding upper bound (UB) is

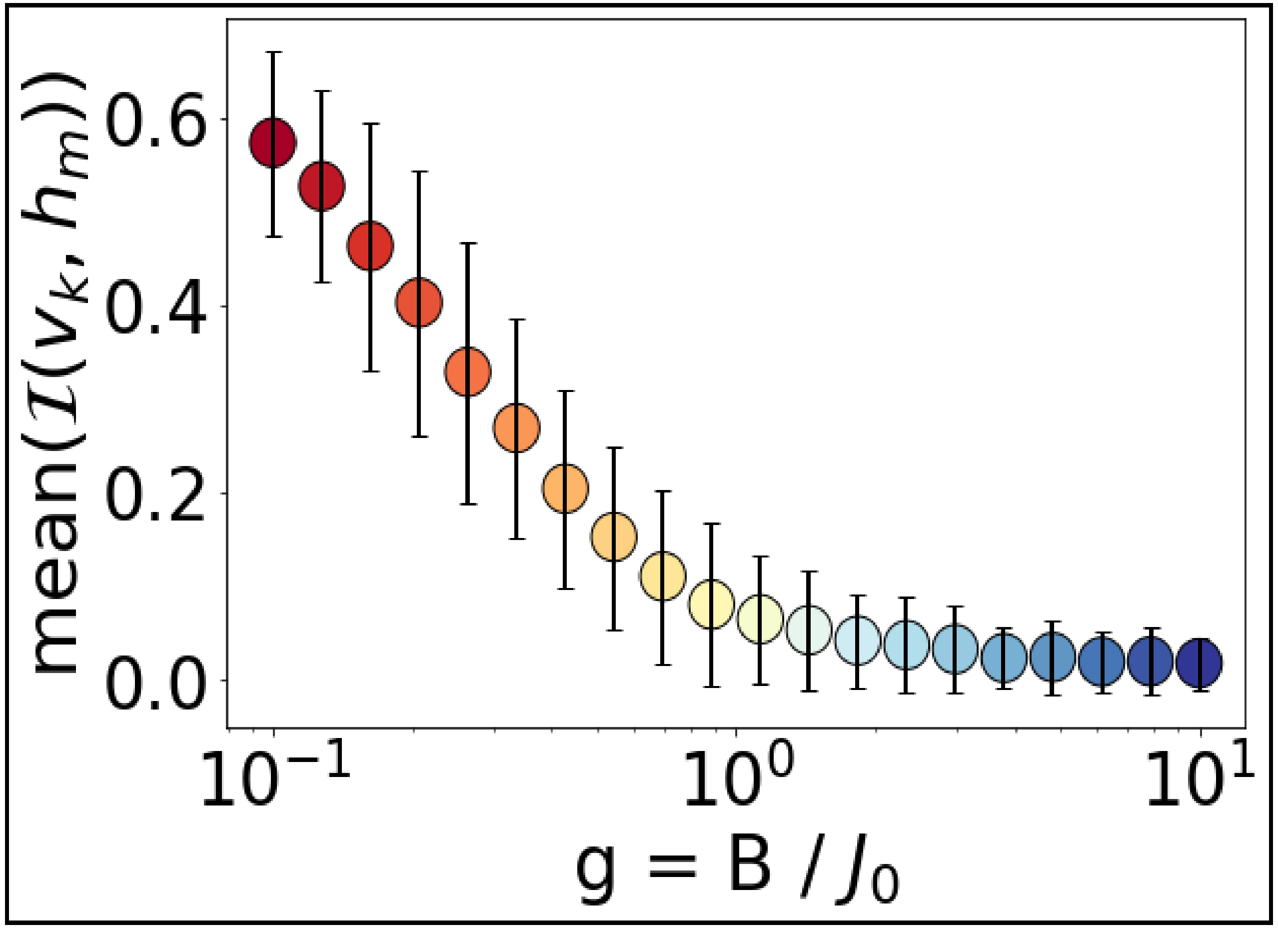
$$UB = (\frac{1}{2} + \frac{\sqrt{1-\eta(\vec{X})}}{2})ln(\frac{1}{2} + \frac{\sqrt{1-\eta(\vec{X})}}{2})$$
$$-(\frac{1}{2} - \frac{\sqrt{1-\eta(\vec{X})}}{2})ln(\frac{1}{2} - \frac{\sqrt{1-\eta(\vec{X})}}{2})$$

with $\eta(\vec{X}) = Im(C_{km}^{km}(\vec{X}, t = \frac{\pi}{2W_{m}^{k}}))$

$$LB \leq \mathcal{I}(v_k, h_m) \leq UB$$

$$\eta(\vec{X}) = Im(C_{km}^{km}(\vec{X}, t = \frac{\pi}{2W_m^k}))$$





$$H = B\sum_i^N \sigma_i^x + J_0 \sum_i^{N-1} \sigma_i^z \sigma_{i+1}^z$$

SYSTEMS TO STUDY (DRIVERS)

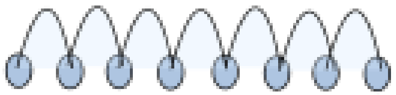
Our results will be primarily focused on two different spin-systems generically described as

$$H = -B\sum_{i_d}^{N}\sigma^x(i_d) - \sum_{i_d j_d} J_{i_d j_d}\sigma^z(i_d)\sigma^z(j_d)$$

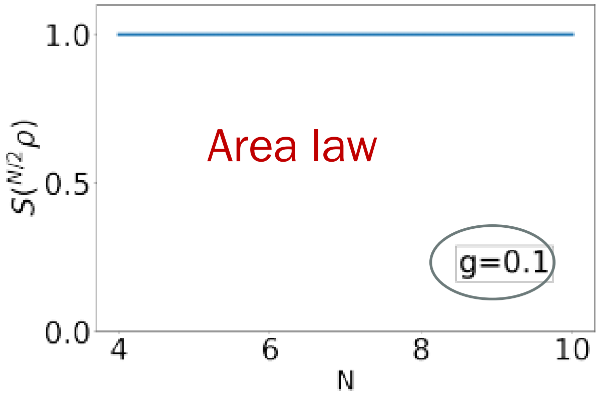
$g = B/J_0$

SYSTEM A

TFIM



$$J_{i_d j_d} = \begin{cases} J_0, & \text{if } if\ j_d = i_d \pm 1\ \ \forall\ i_d \\ 0 & \text{otherwise} \end{cases}$$

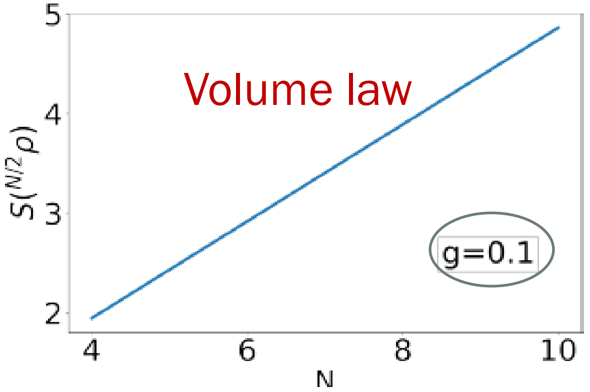


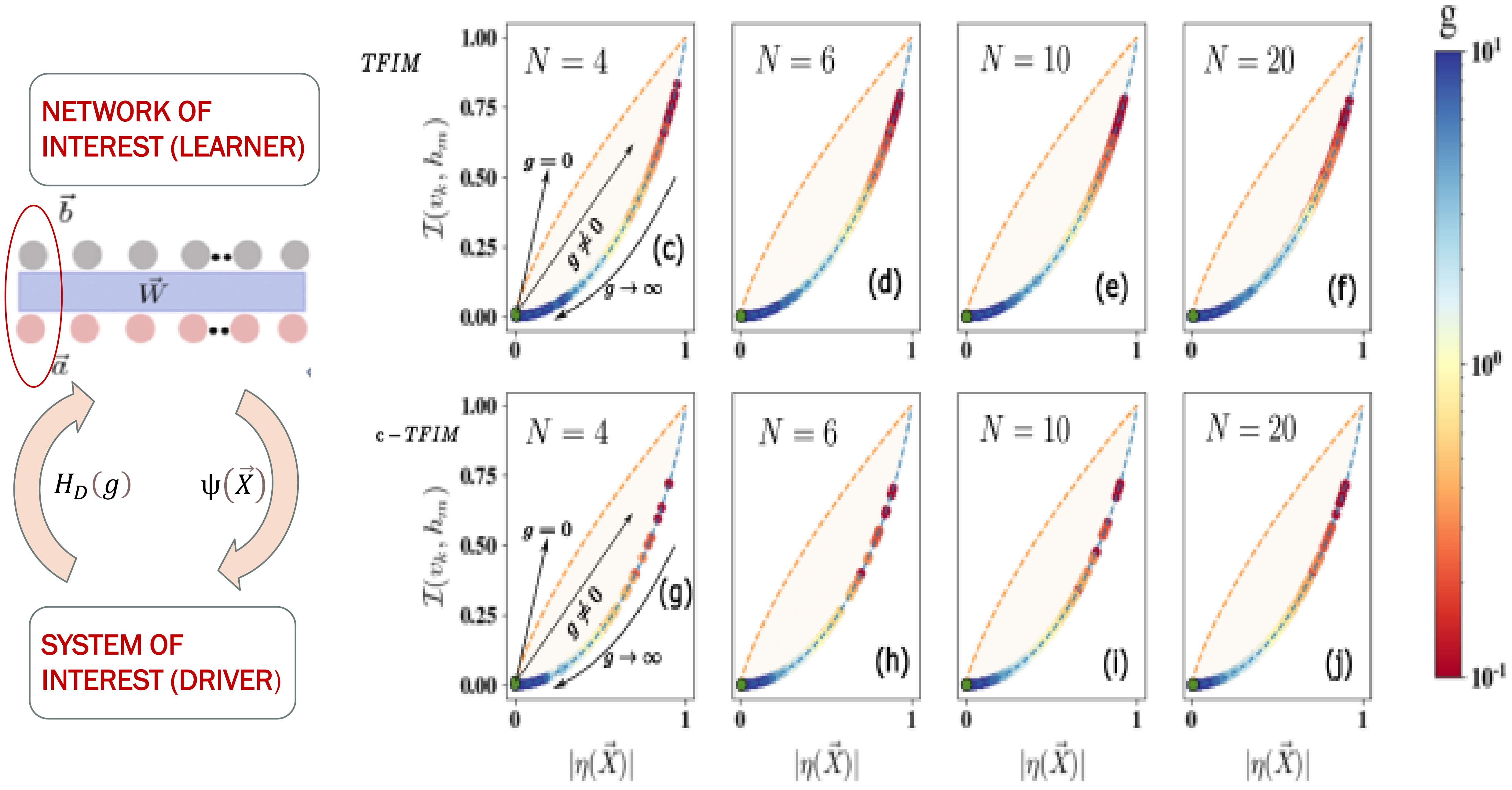
SYSTEM B

$c-TFIM$

$$J_{i_d j_d} = \begin{cases} J_0, & \text{if } i_d = \frac{N}{2} - (q-1), \\ & j_d = \frac{N}{2} + q \\ 0 & \text{otherwise} \end{cases}$$

with $\forall\ q \in [1, \frac{N}{2}]$



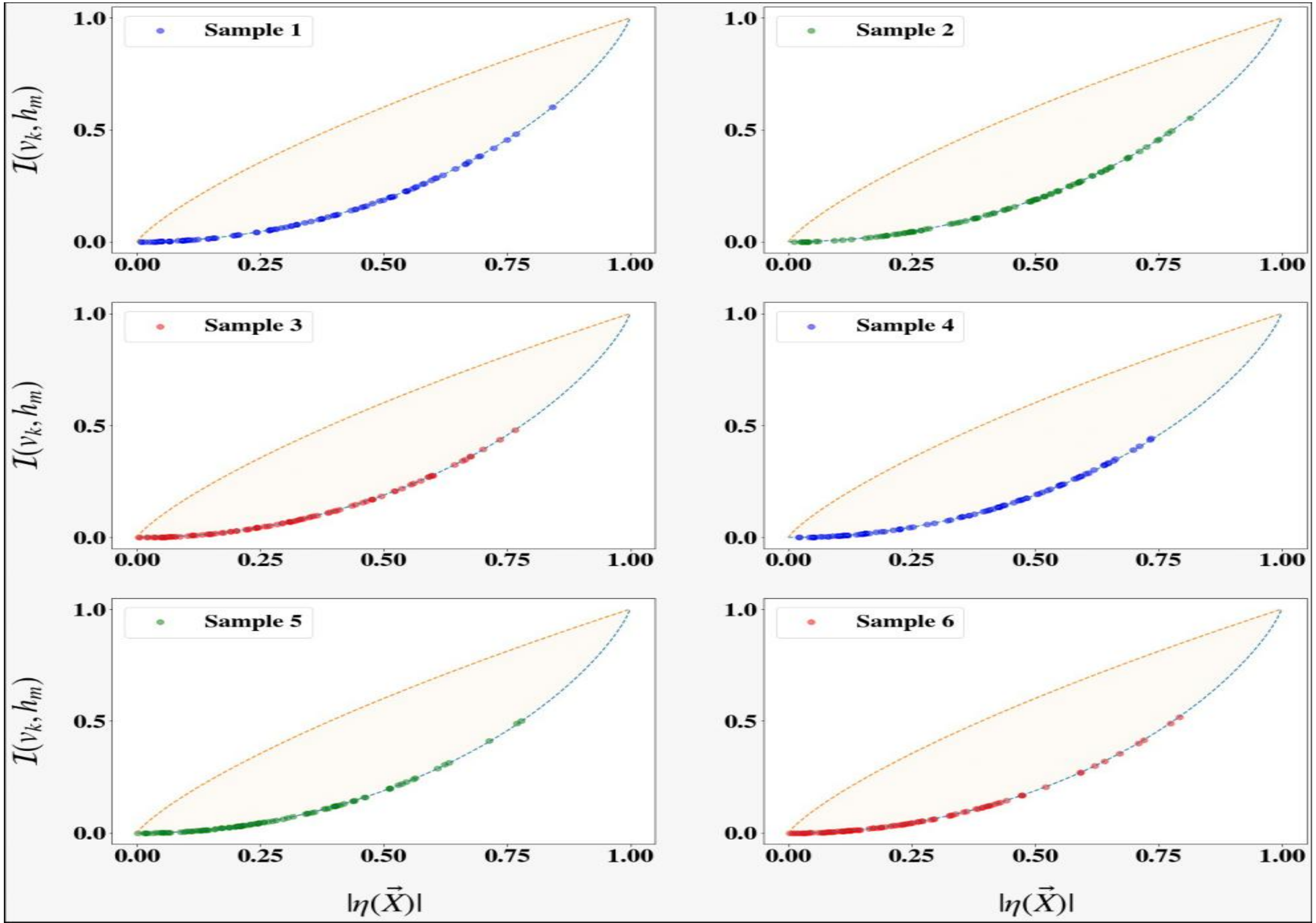


NETWORK OF INTEREST (LEARNER)
$\vec{b}$
$\vec{W}$
$\vec{\alpha}$
$H_D(g)$
$\psi(\vec{X})$
SYSTEM OF INTEREST (DRIVER)
TFIM
c − TFIM
$\mathcal{I}(v_k, h_m)$
$|\eta(\vec{X})|$
N = 4
N = 6
N = 10
N = 20
g = 0
g ≠ 0
g → ∞
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
g

Sherrington-Kirkpatrick model

$$H_D = -B\sum_{i_d}^{N}\sigma^x(i_d) - \sum_{i_d j_d} J_{i_d j_d}\sigma^z(i_d)\sigma^z(j_d)$$

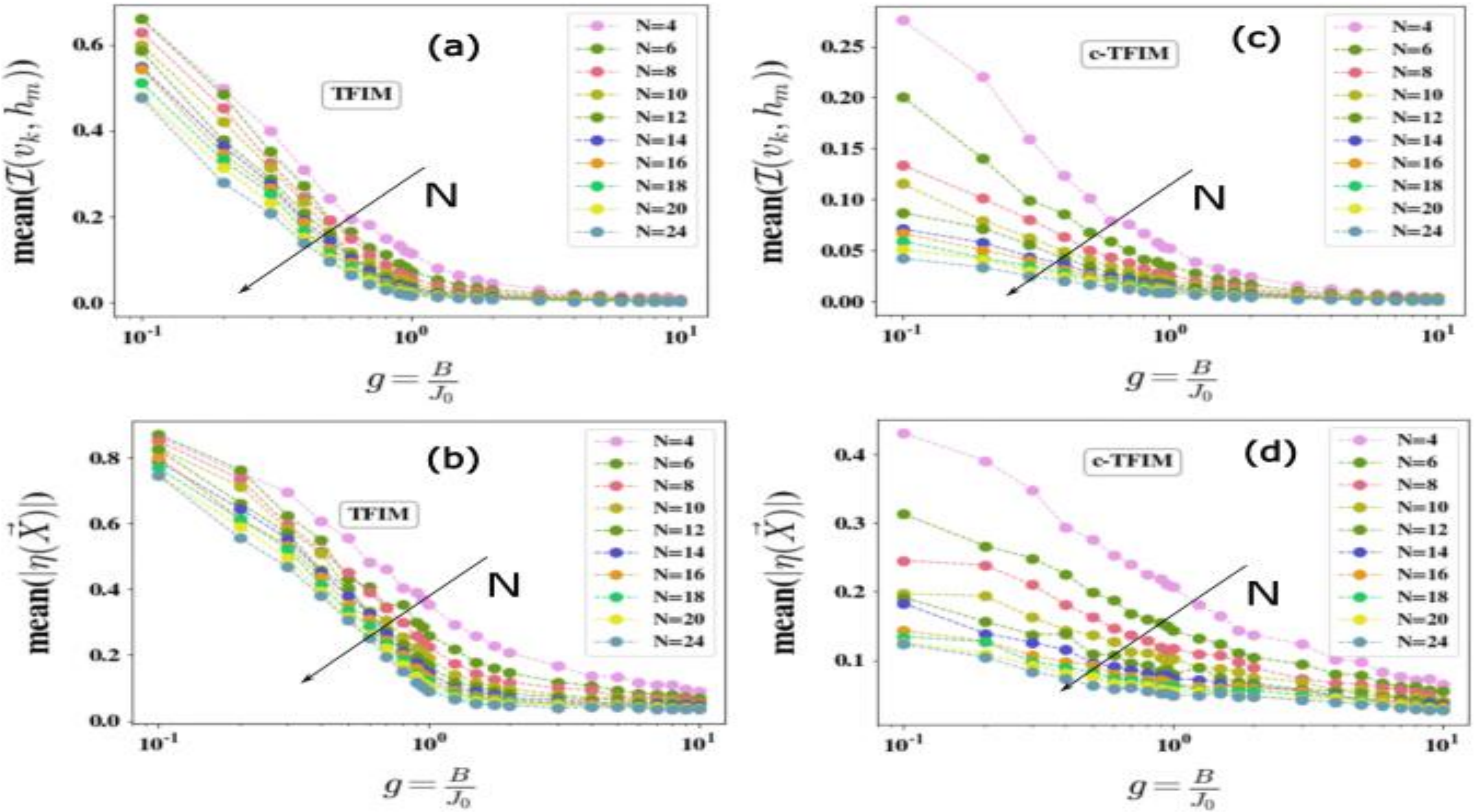
$J_{i_d j_d}$ are sampled from $\mathcal{N}(0,1)$.



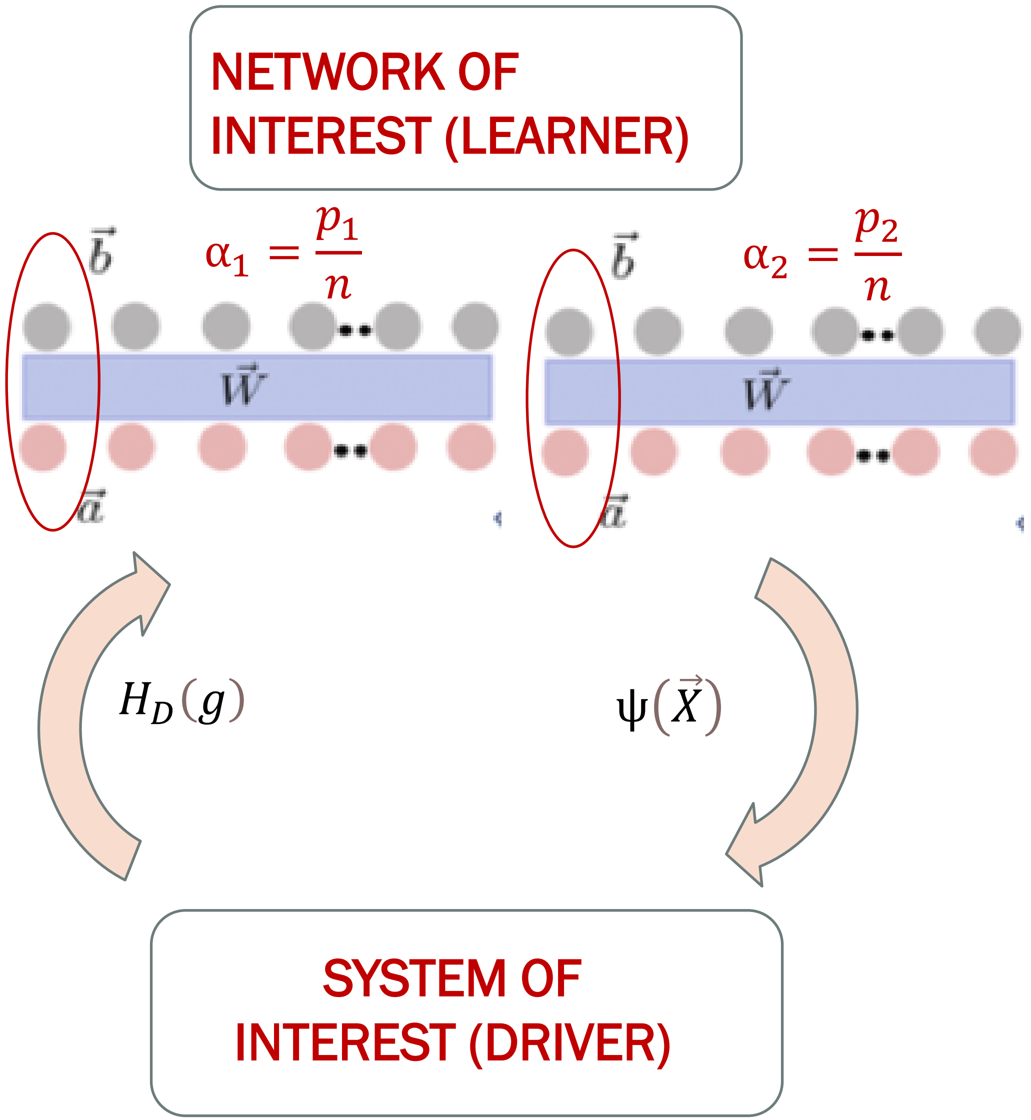
$$H = -B\sum_{i_d}^{N}\sigma^x(i_d) - \sum_{i_d j_d} J_{i_d j_d}\sigma^z(i_d)\sigma^z(j_d)$$

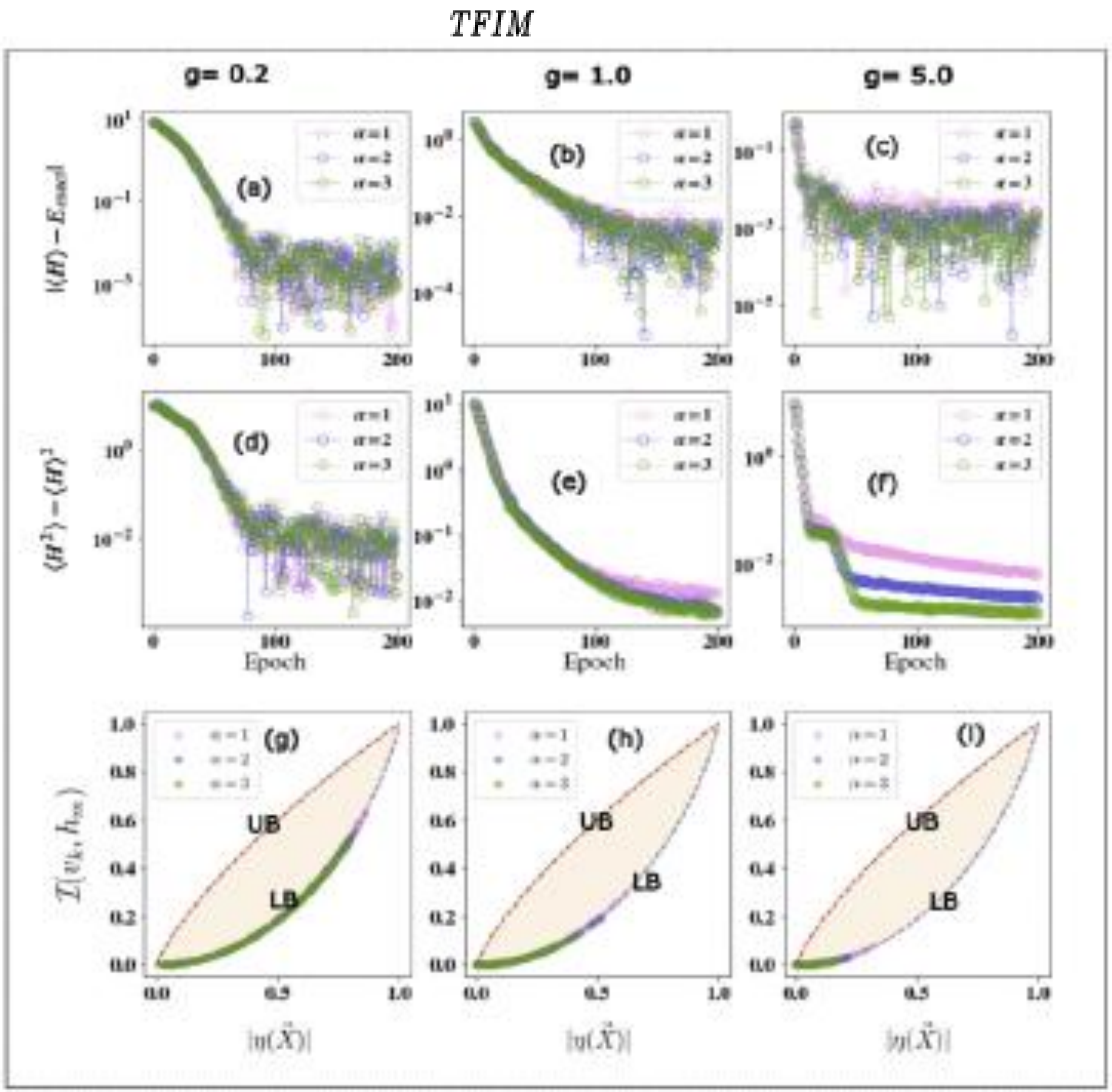
TFIM

c−TFIM



EFFECT OF HIDDEN NODE DENSITY





Summary of OTOC Work

- We analytically illustrate that the imaginary components of OTOC can be related to mutual information
- We rigorously establish the mathematical bounds on such quantities respected by the dynamical evolution during training
- OTOC offers important insights into the training dynamics by unraveling how quantum information is scrambled through such a network introducing correlation among its constituent sub-systems

PHYSICAL REVIEW RESEARCH 5, 013146 (2023)

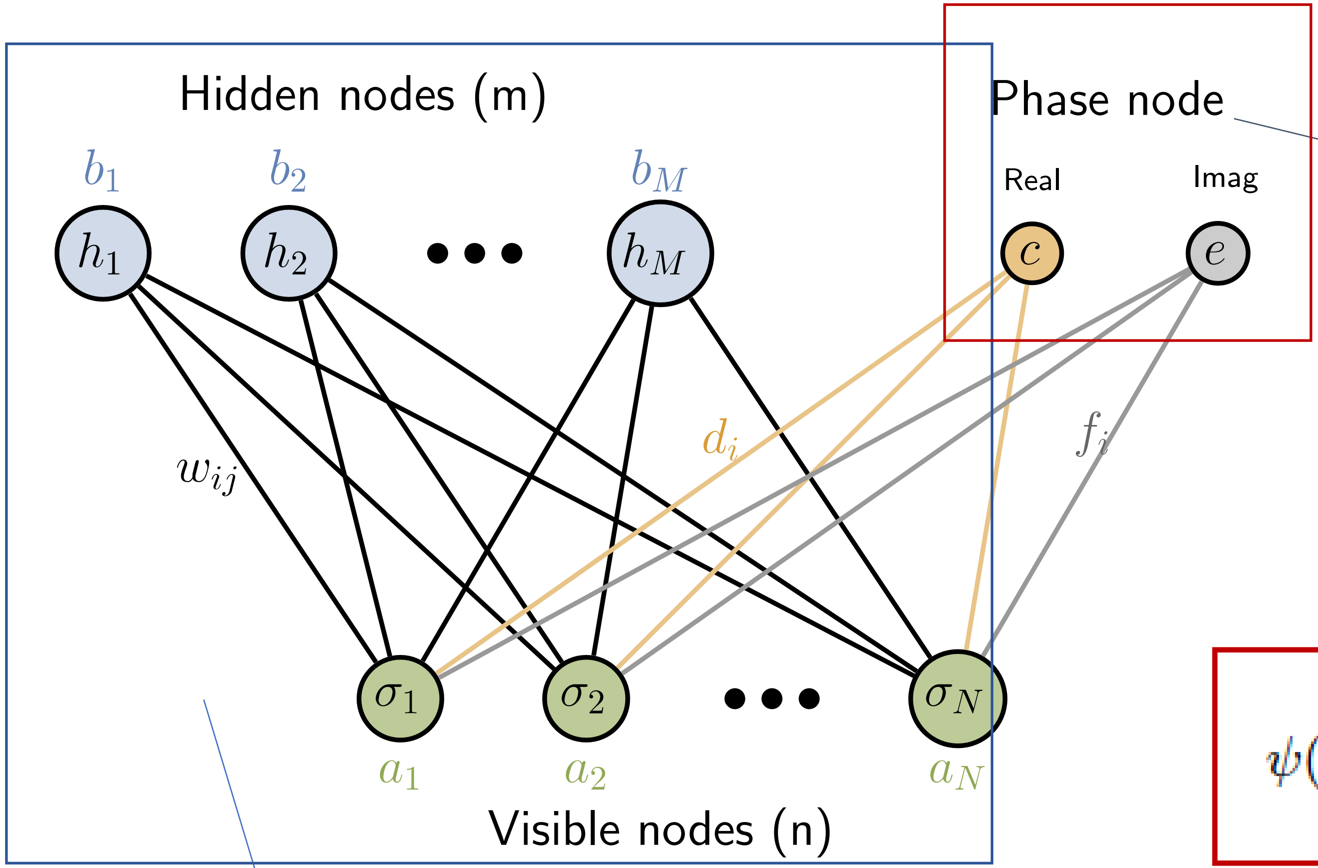
Future Work

- Is it possible to leverage power of quantum devices into the workflow? Do we see any advantage there?

- Can we design such a workflow for an arbitrary NQS given their structural diversity ?

- Can we make the NQS encode both amplitude and phase of the target and reduce storage keeping runtime and circuit requirements polynomial and do better than previous algos ?

Quantum Machine Learning Algorithm



$$s(\vec{d}, \vec{f}, c, e, \vec{\sigma}) = \tanh\left[(c + \sum_i d_i\sigma_i) + i(e + \sum_i f_i\sigma_i)\right]$$

$$\psi(\vec{X}) = \sum_\sigma \sqrt{\tilde{p}(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma}) s(\vec{d}, \vec{f}, c, e, \vec{\sigma})} |\sigma_1\sigma_2 \ldots \sigma_n\rangle$$

$$P(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma}, \vec{h}) = \frac{e^{\sum_i a_i\sigma_i + \sum_j b_j h_j + \sum_{ij} w_{ij}\sigma_i h_j}}{\sum_{\{\sigma h\}} e^{\sum_i a_i\sigma_i + \sum_j b_j h_j + \sum_{ij} w_{ij}\sigma_i h_j}}$$

$$\tilde{p}(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma}) = \sum_h P(\vec{a}, \vec{b}, \vec{W}, \vec{\sigma}, \vec{h})$$

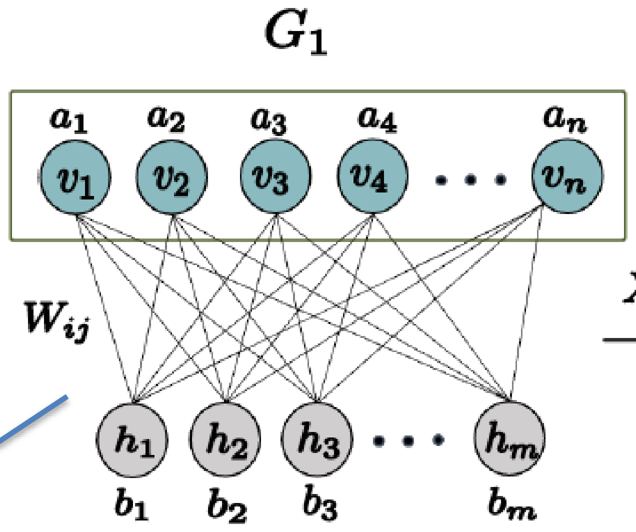
Introducing surrogate networks- Generalizability to arbitrary NQS

Concrete example with a given Neural Quantum State (NQS)

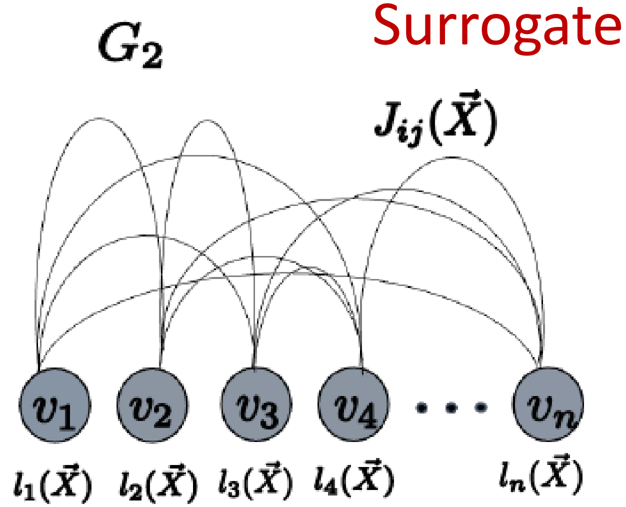
Original NQS- RBM

$\rho_v^{v}(\vec{X}) \propto \kappa(\vec{v}, \vec{X})\phi_{\vec{v}}(\vec{l}(\vec{X}), \vec{J}(\vec{X}))$

Surrogate for k=2



$\vec{X} = (\vec{a}, \vec{b}, \vec{W})$



$$\rho_{v'}^{v}(\vec{X}) \propto e^{-\beta \sum_i a_i v_i + \sum_i a_i^* v_i'} \prod_{j=1}^{m} \Gamma_j(\vec{v}, \vec{b}, \vec{W})\, \Gamma_j(\vec{v}', \vec{b}^*, \vec{W}^*)$$

$$\Gamma_j(\vec{v}, \vec{b}, \vec{W}) = \mathrm{Cosh}(\beta \mathrm{b_j} + \beta \sum_{i=1}^{n} \mathrm{W_{ij} v_i}).$$

$$\phi_{\vec{v}}(\vec{l}(\vec{X}), \vec{J}(\vec{X})) \propto e^{-\beta \sum_i l_i(\vec{X}) v_i + \sum_{ij} J_{ij}(\vec{X}) v_i v_j}$$

Polynomially efficient quantum enabled variational Monte Carlo for training neural-network quantum states for physico-chemical applications

Manas Sajjan†,[1,2] Vinit Singh†,[1,2] and Sabre Kais[2,*]

[1]*Department of Chemistry, Purdue University, West Lafayette, IN 47907*
[2]*Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27606*

arXiv:2412.12398v1

What do we achieve by defining a new surrogate ?

- We can map an arbitrary distribution associated with any NQS architecture to a single layer of neurons (surrogate) with certain connectivity pattern. This lends generalizability

- The surrogate distribution is easy to sample. We can use the form of the surrogate distribution to define a quantum enabled workflow for any NQS

- We can achieve variance reduction in the estimated observable using a proper choice of prefactor

$$\rho_v^v(\vec{X}) \propto \kappa(\vec{v}, \vec{X})\phi_{\vec{v}}(\vec{l}(\vec{X}), \vec{J}(\vec{X}))$$

Original NQS

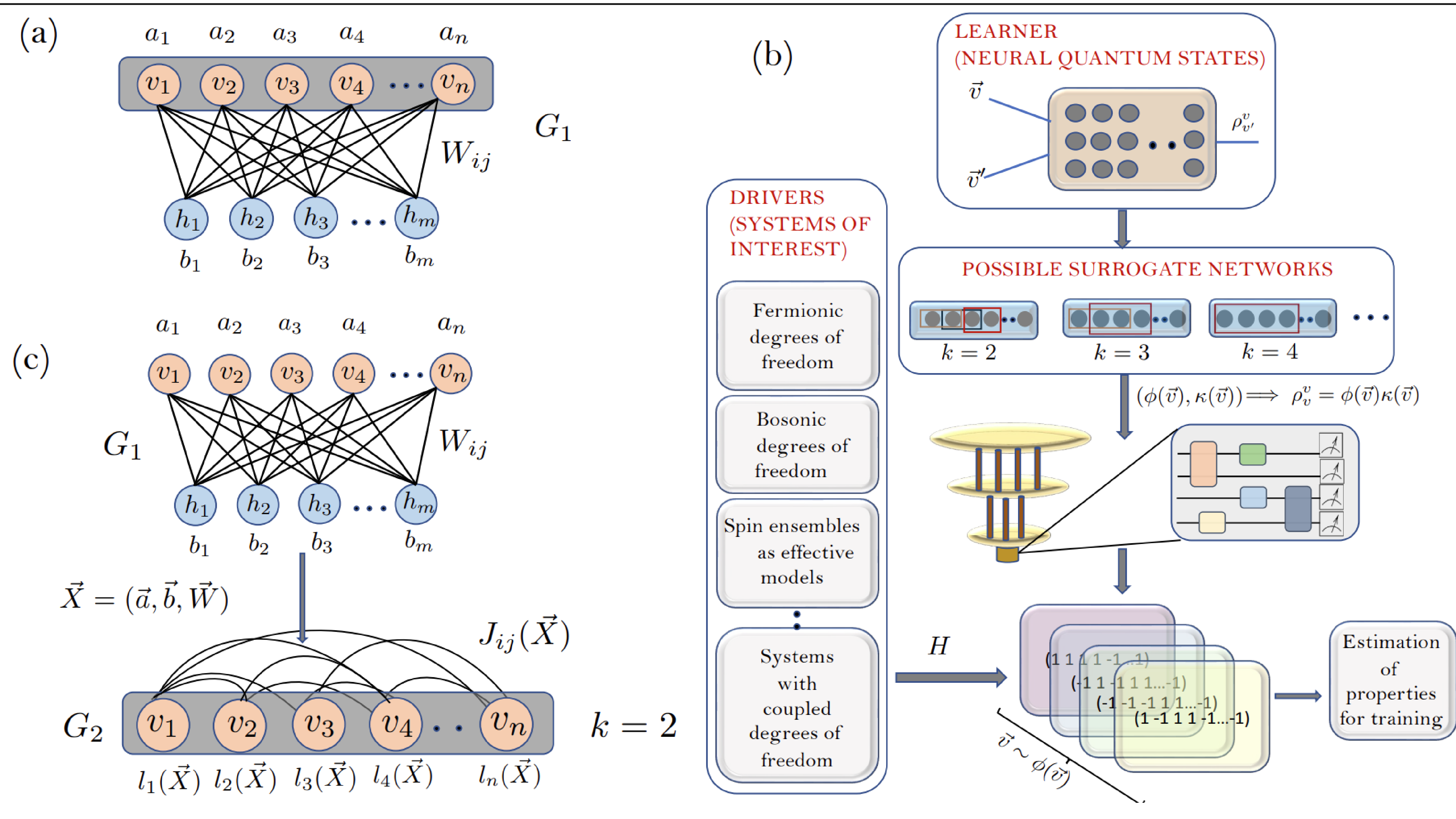
$$\langle H\rangle(\vec{X}) = \frac{Tr(\rho(\vec{X})H)}{Tr(\rho(\vec{X}))} = \sum_{\vec{v}} \frac{\rho_v^v(\vec{X})E_{loc}(\vec{v}, \vec{X})}{\sum_{\vec{v}} \rho_v^v(\vec{X})}$$

Surrogate NQS

$$\langle H\rangle(\vec{X}) \approx \frac{\sum_{\vec{v}} \kappa(\vec{v}, \vec{X})\phi_{\vec{v}}(\vec{l}(\vec{X}), \vec{J}(\vec{X}))E_{loc}(\vec{v}, \vec{X})}{\sum_{\vec{v}} \kappa(\vec{v}, \vec{X})\phi_{\vec{v}}(\vec{l}(\vec{X}), \vec{J}(\vec{X}))}$$

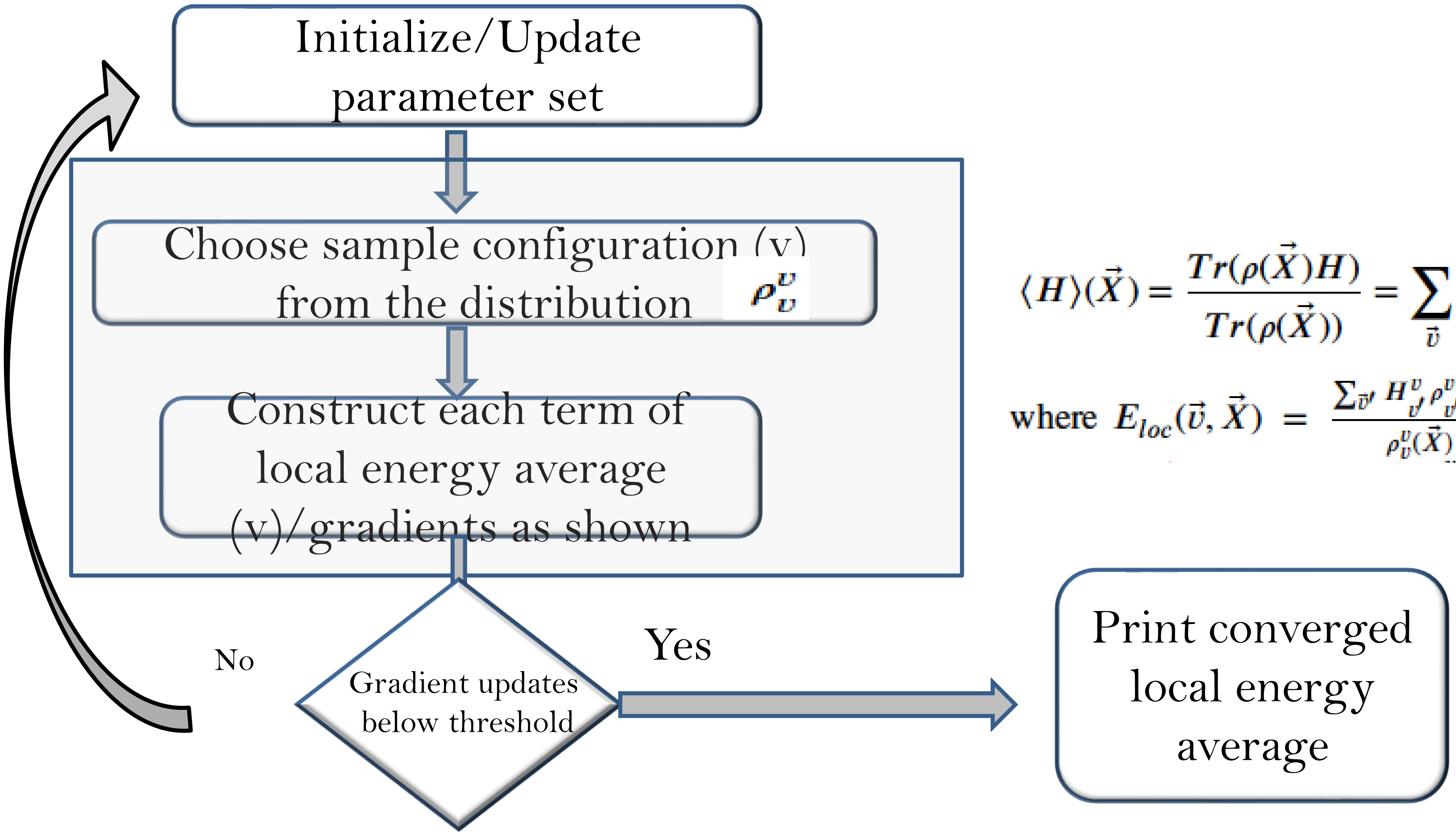
$$\downarrow$$

$$\mu_{\langle H\rangle(\vec{X})} = \frac{\sum_{\vec{v}\sim\phi} \kappa(\vec{v}, \vec{X})E_{loc}(\vec{v}, \vec{X})}{\sum_{\vec{v}\sim\phi} \kappa(\vec{v}, \vec{X})}$$



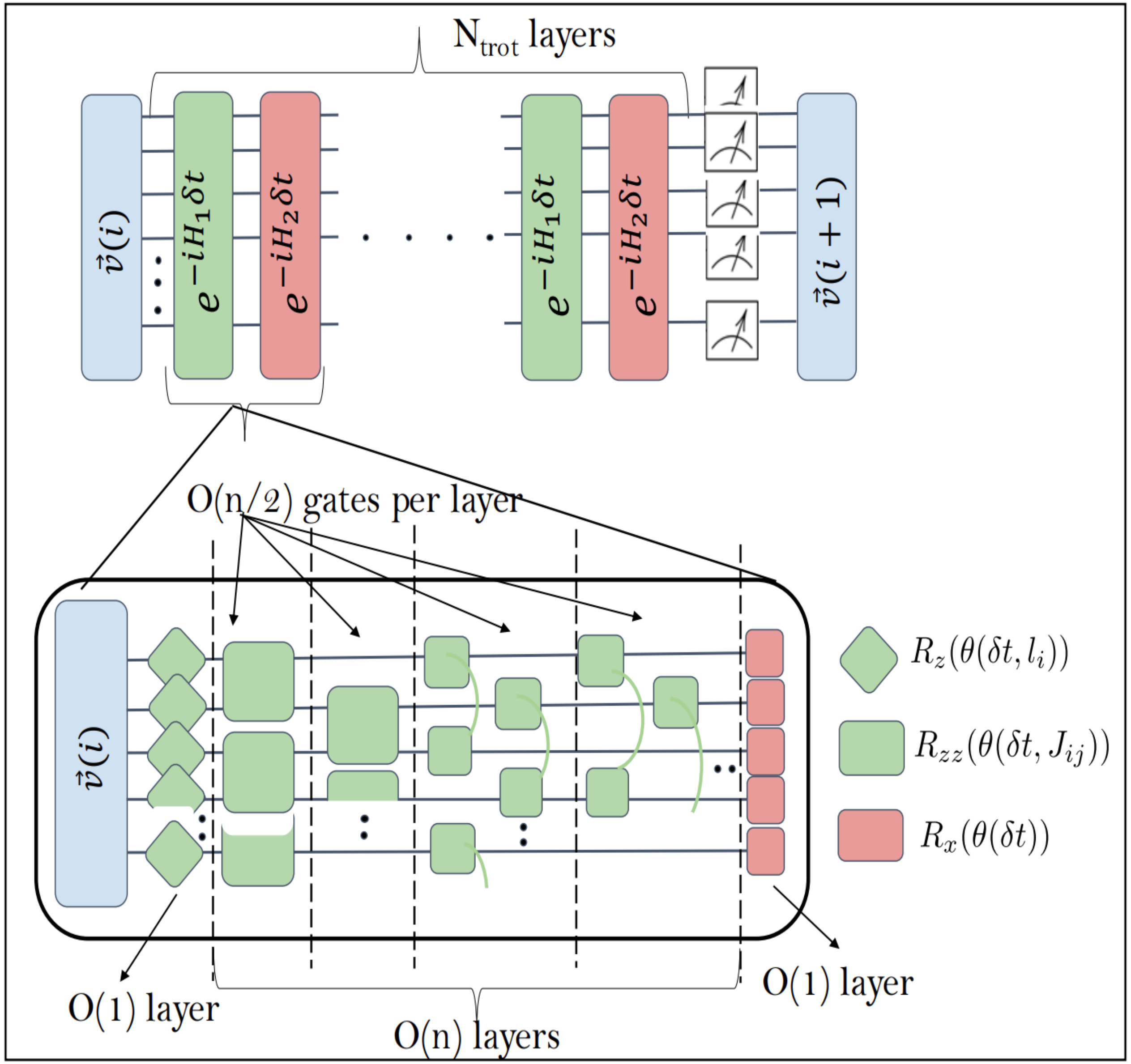
(a)
a_1 a_2 a_3 a_4 a_n
v_1 v_2 v_3 v_4 ... v_n
G_1
W_{ij}
h_1 h_2 h_3 ... h_m
b_1 b_2 b_3 b_m
(b)
LEARNER (NEURAL QUANTUM STATES)
$\vec{v}$
$\vec{v}'$
$\rho^v_{v'}$
DRIVERS (SYSTEMS OF INTEREST)
Fermionic degrees of freedom
Bosonic degrees of freedom
Spin ensembles as effective models
Systems with coupled degrees of freedom
POSSIBLE SURROGATE NETWORKS
$k = 2$ $k = 3$ $k = 4$
$(\phi(\vec{v}), \kappa(\vec{v})) \Longrightarrow \rho^v_v = \phi(\vec{v})\kappa(\vec{v})$
H
(1 1 1 1 -1...1)
(-1 1 -1 1 1...-1)
(-1 -1 -1 1 1...-1)
(1 -1 1 1 -1...-1)
$\vec{v} \sim \phi(\vec{v})$
Estimation of properties for training
(c)
a_1 a_2 a_3 a_4 a_n
v_1 v_2 v_3 v_4 ... v_n
G_1
W_{ij}
h_1 h_2 h_3 ... h_m
b_1 b_2 b_3 b_m
$\vec{X} = (\vec{a}, \vec{b}, \vec{W})$
$J_{ij}(\vec{X})$
G_2
v_1 v_2 v_3 v_4 .. v_n
$k = 2$
$l_1(\vec{X})$ $l_2(\vec{X})$ $l_3(\vec{X})$ $l_4(\vec{X})$ $l_n(\vec{X})$

Usual workflow for training to learn energy eigenstates



$$\langle H\rangle(\vec{X}) = \frac{Tr(\rho(\vec{X})H)}{Tr(\rho(\vec{X}))} = \sum_{\vec{v}} \frac{\rho^{v}_{v}(\vec{X})E_{loc}(\vec{v},\vec{X})}{\sum_{\vec{v}} \rho^{v}_{v}(\vec{X})}$$

where $E_{loc}(\vec{v},\vec{X}) = \frac{\sum_{\vec{v}'} H^{v}_{v'}\rho^{v}_{v'}(\vec{X})}{\rho^{v}_{v}(\vec{X})}$ is the local energy.



# qubits	Gate depth	# measurement
O(n)	O(n)	O(Ns) (samples drawn)

Our algorithm requires Rzz and single qubit gates to be implemented. Since each Rzz has 2 CNOT, so CNOT depth is O(2n)

Comparison with previous known algorithm

# qubits	Gate depth	# measurement
O(m+n)	O(n^2)	O(n^4)

Best known quantum variational algorithms for chemistry – like those using UCCSD ansatz has depth of O(n^4) and O(n^4) parameters yet achieves comparable accuracy. For other ansatz types like HEA, HVA the depth is problem dependent and they are often less expressive .


JACS

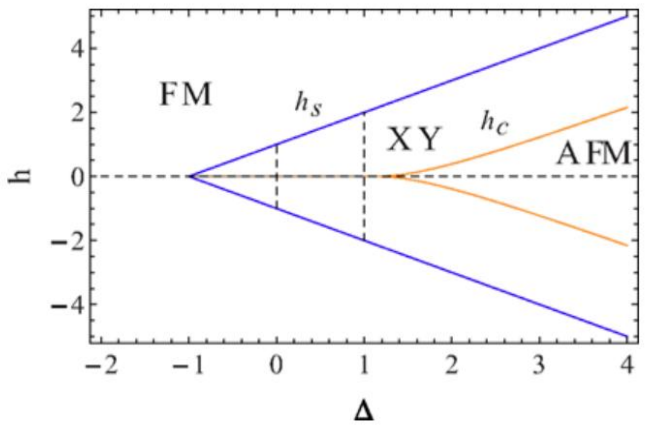
Quantum Machine-Learning for Eigenstate Filtration in Two-Dimensional Materials

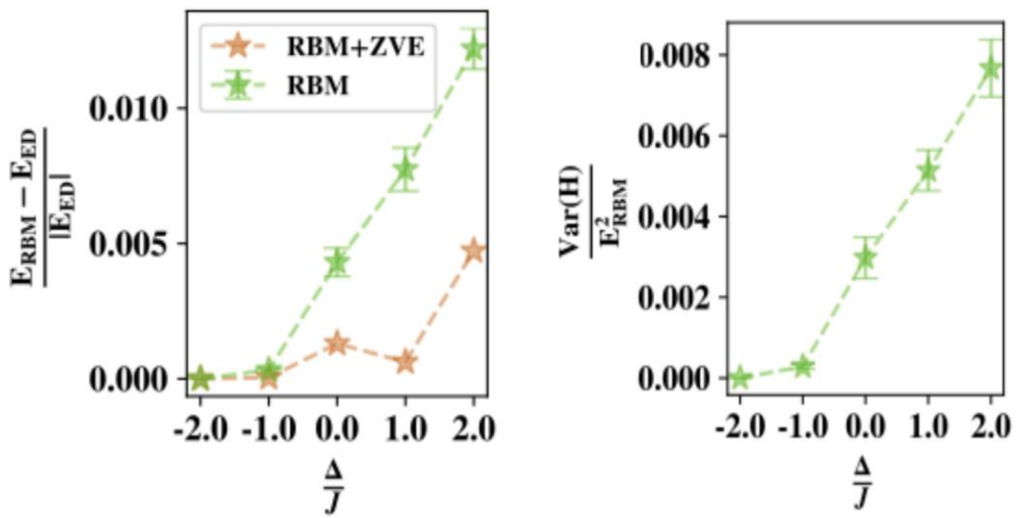
Manas Sajjan, Shree Hari Sureshbabu, and Sabre Kais*

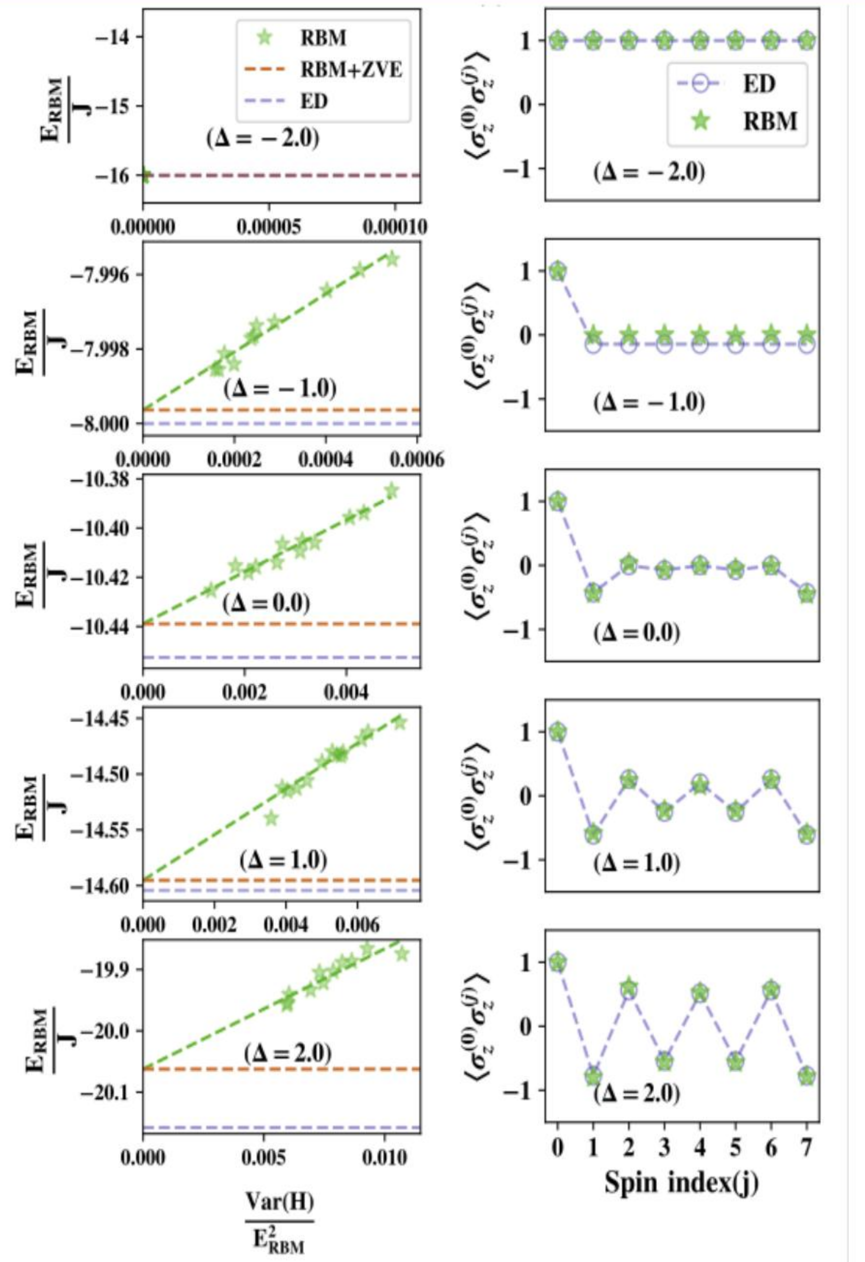
Cite This: J. Am. Chem. Soc. 2021, 143, 18426−18445

Applications (Ground state learning in XXZ)

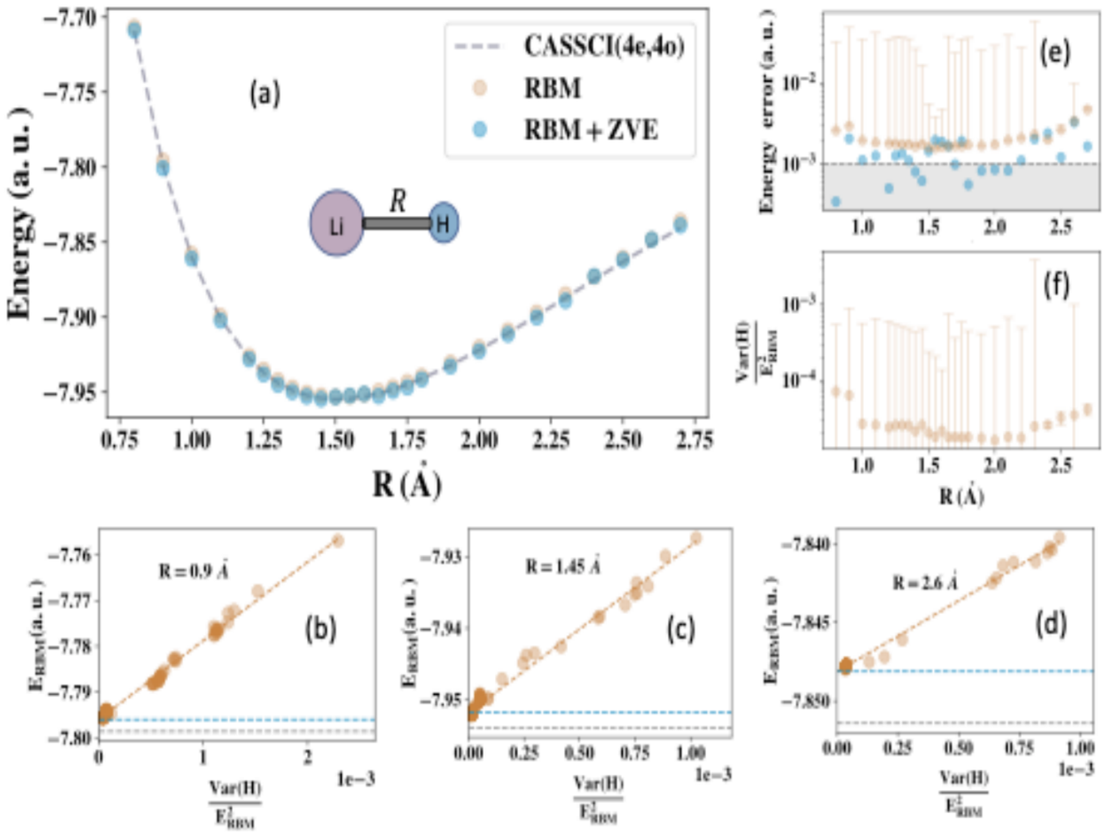
$$H = J\sum_{i=1}^{N}[\sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y + \Delta\sigma_i^z\sigma_{i+1}^z] - 2h\sum_{i=1}^{N}\sigma_i^z$$

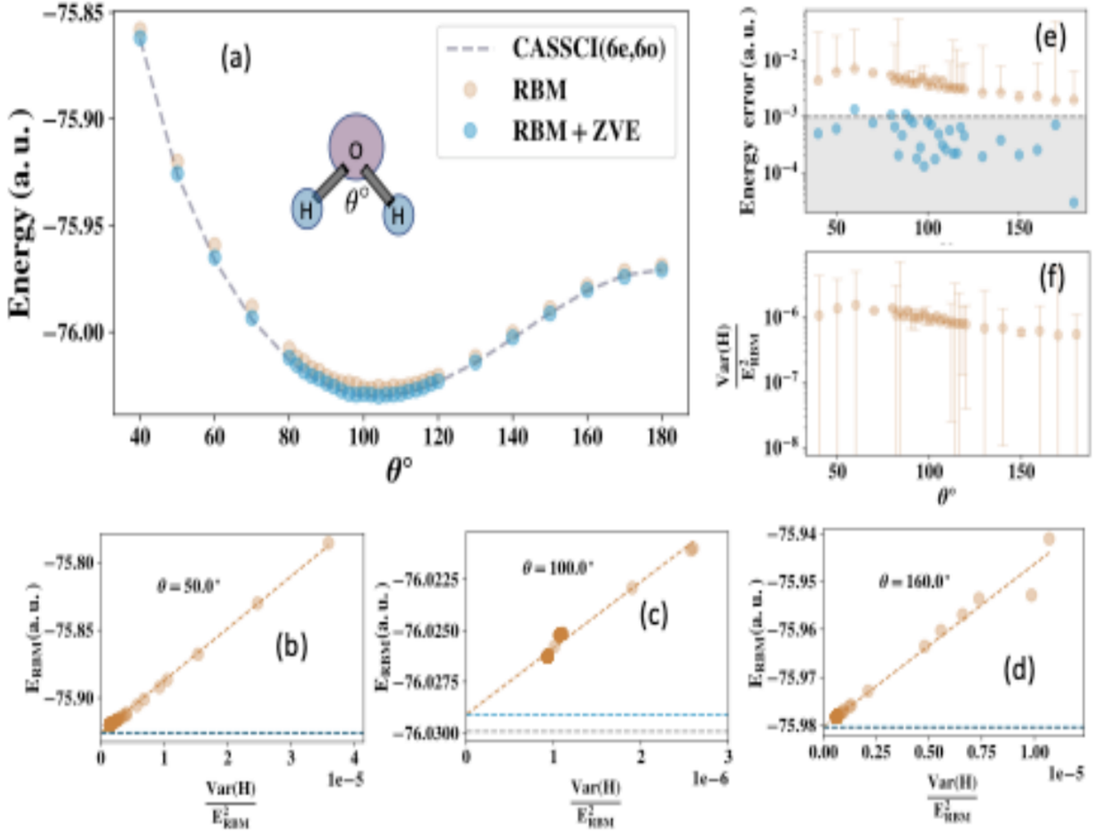


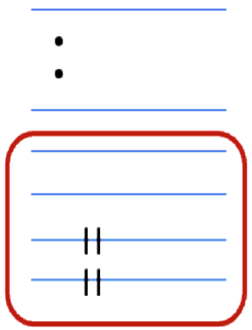




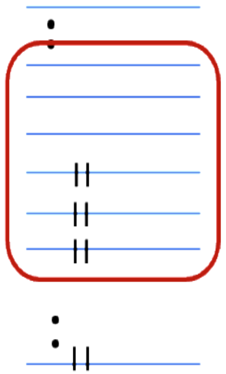
Applications (Ground potential energy surface learning in molecular systems)



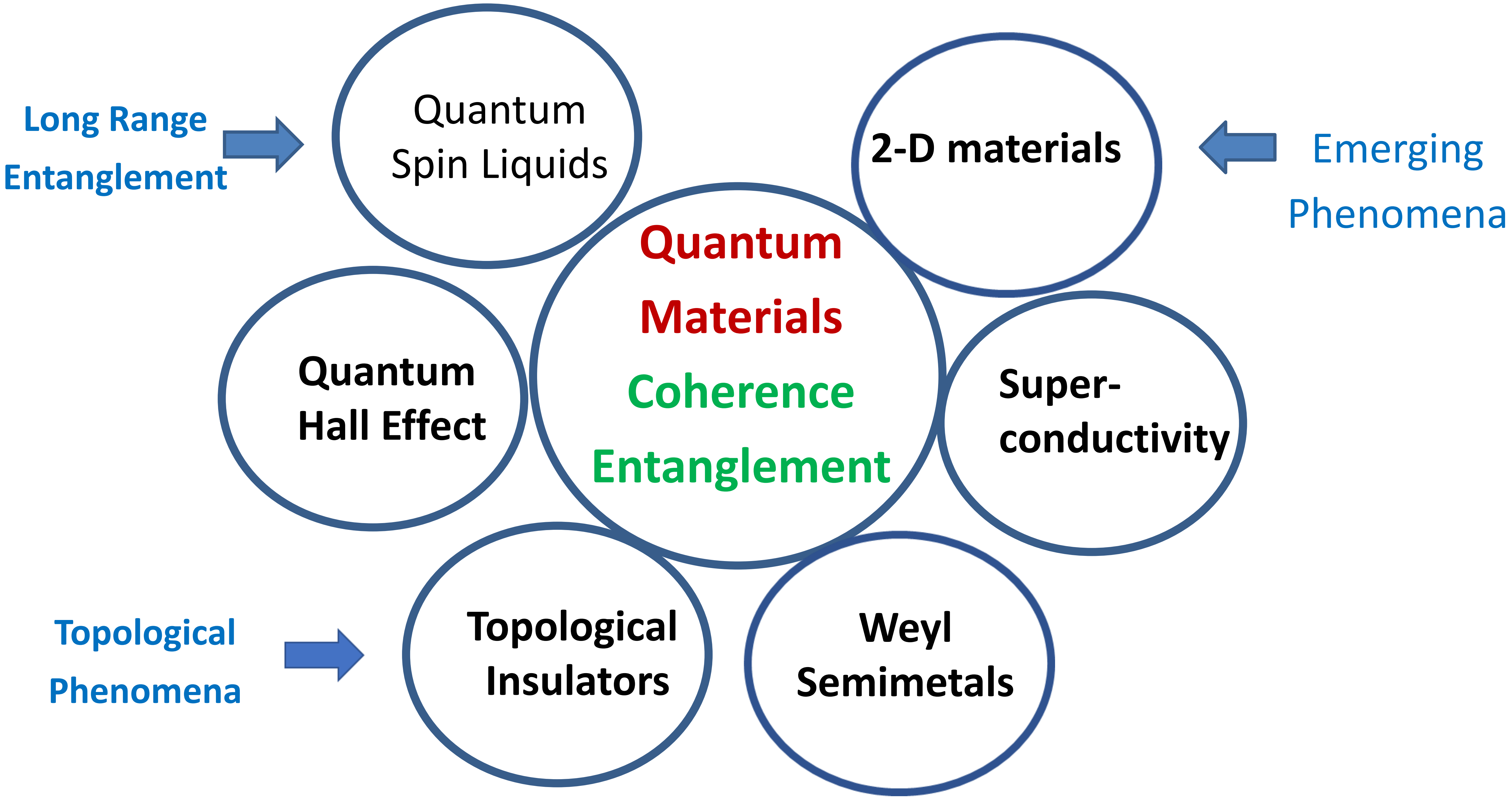




n=8



n=12

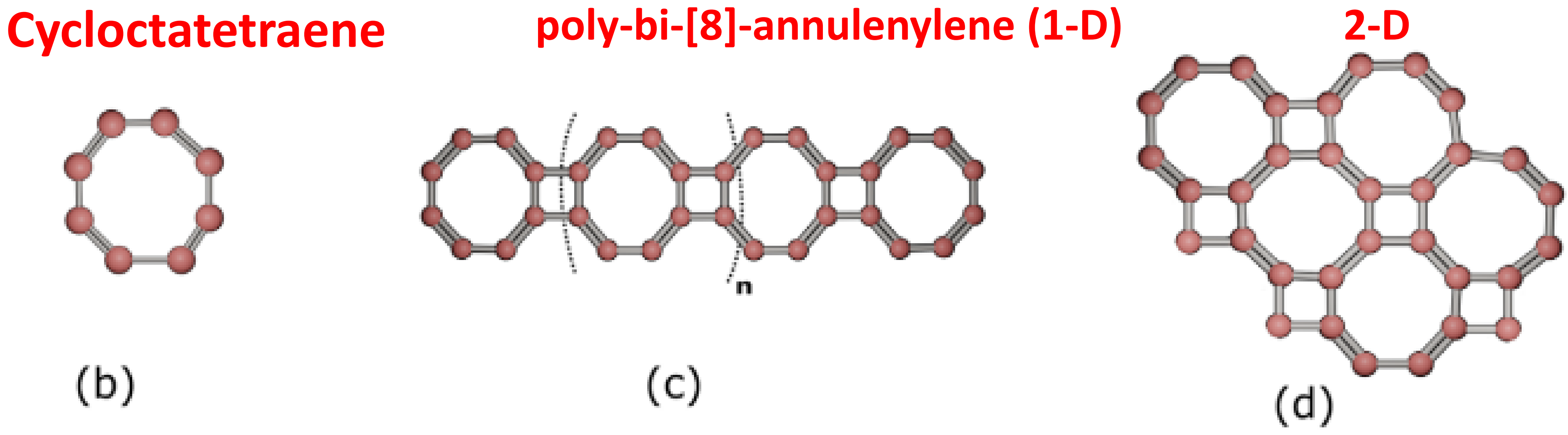

Long Range
Entanglement
Quantum
Spin Liquids
2-D materials
Emerging
Phenomena
Quantum
Materials
Coherence
Entanglement
Quantum
Hall Effect
Super-
conductivity
Topological
Phenomena
Topological
Insulators
Weyl
Semimetals

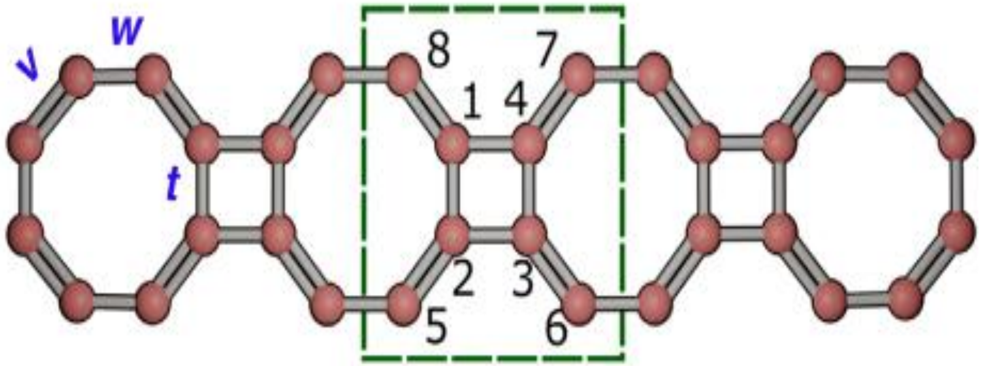
Received: 8 June 2023 | Revised: 4 July 2023 | Accepted: 5 July 2023
DOI: 10.1002/ntls.20230015
Natural Sciences

RESEARCH ARTICLE

Foray into the topology of poly-bi-[8]-annulenylene

Varadharajan Muruganandam[1,2] | Manas Sajjan[2,3] | Sabre Kais[1,2,3]

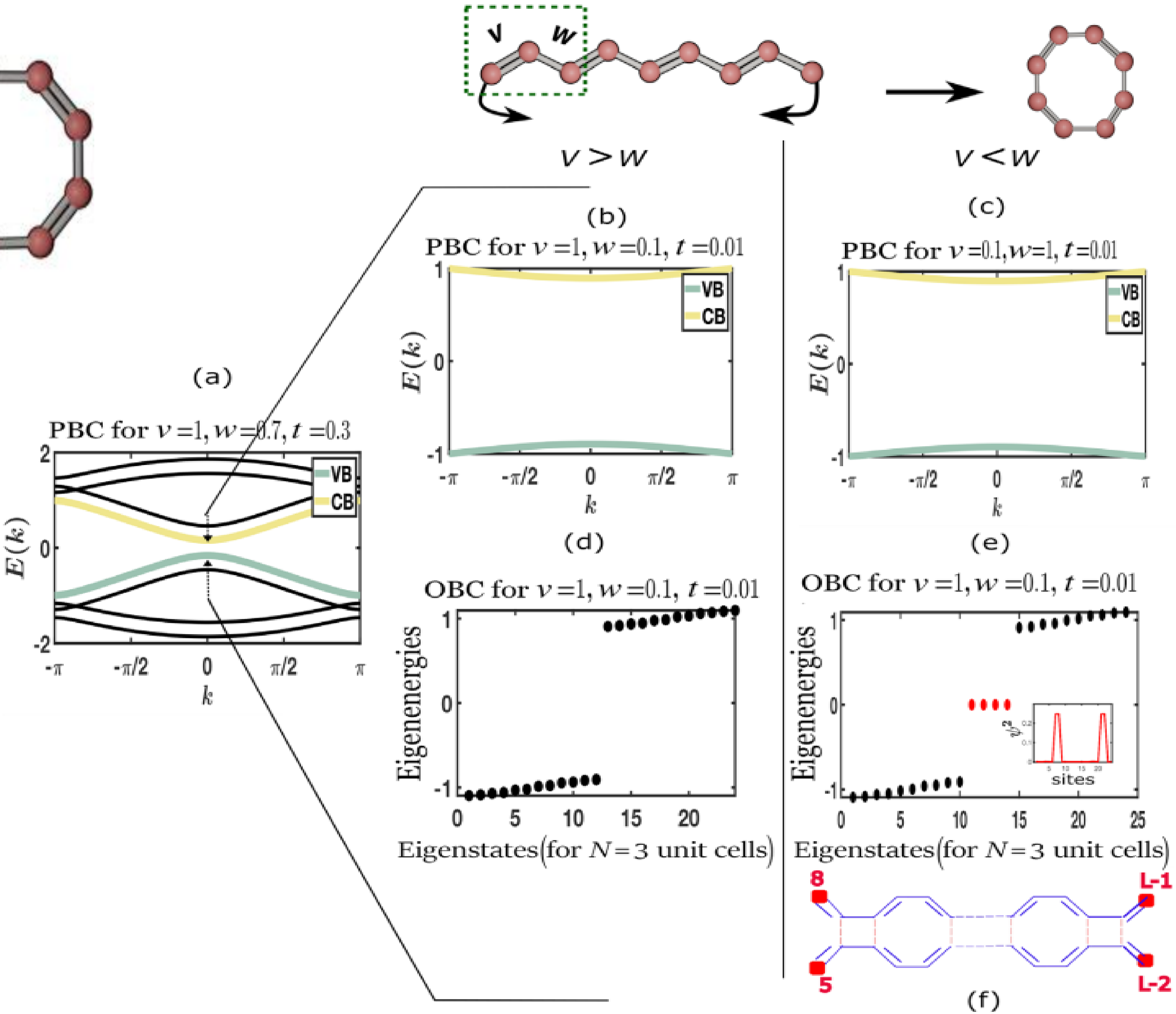




Poly-bi-[8]-annulenylene in one dimension

For poly-bi-[8]-annulenylene, the unit cell marked in green has eight sites. The Hamiltonian of the model is as follows:

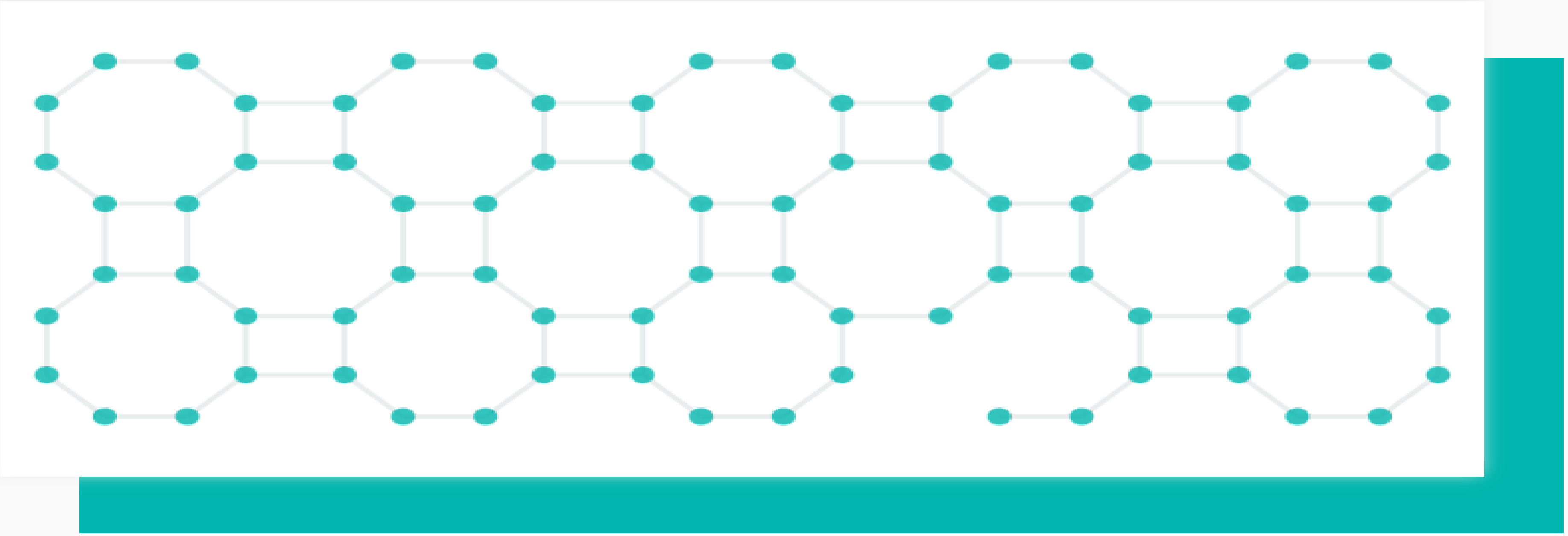
$$H_{PO[8]A} = \sum_{r=1}^{N}[-v(c^{\dagger}_{r,1}c_{r,8} + c^{\dagger}_{r,4}c_{r,7} + c^{\dagger}_{r,2}c_{r,5}) - w(c^{\dagger}_{r,1}c_{r,4} + c^{\dagger}_{r,2}c_{r,3}) - t(c^{\dagger}_{r,1}c_{r,2} + c^{\dagger}_{r,4}c_{r,3})] + h.c.$$



Cheryl D, Stevenson JPD. Cyclooctatetraene-based cathode for electrochemical cells. 2010. US Patent US20100288628A1.

Rigetti Systems

Aspen-M-3 Quantum Processor



Conclusions

- Restricted Boltzmann Machine (RBM) can be used to perform electronic structure calculations:
  H_2, H_2O, LiH, h-BN, graphene, Molybdenum disulfide(MoS_2) and Tungsten disulfide (WS_2)

- We **have quadratic resource requirements (circuit width, circuit depth, parameter count).**

- FSS combined with RBM can be used to implement **quantum phase transitions** on quantum devices.

- We trained the network on various flavors of computation using not only a classical computer, Qasm quantum simulator in Qiskit **but also a real IBMQ machine**

- Our algorithm demonstrated **very high accuracy** when compared to the exact values obtained from direct diagonalization.

- QML is a promising approach for treating **open quantum dynamics problems**

Strikingly, the RBM representation for these states is remarkably efficient, in the sense that the number of nonzero parameters scales only linearly with the system size

NC STATE SUMMER SCHOOL 2026

Quantum Machine Learning for Complex Science and Engineering Problems

NC State University June 1 – June 5, 2026

Explore the Future of Quantum Machine Learning

Join leading experts and emerging researchers for a week-long event focused on quantum machine learning (QML) and hybrid quantum-classical methods for solving real-world problems

Kais Research Group

Quantum Information and Quantum Computing

12 Ph.D. and 5 Post

Chem, Phys, CS, ECE

Funding

Thank You