NEURAL - NETWORK QUANTUM STATES

GIUSEPPE CARLEO

EPFL - SWITZERLAND

USEFUL REFERENCES

GENERAL ML

"PROBABILISTIC MACHINE LEARNING" BOOKS I – II
BY KEVIN MURPHY
$\boxed{\text{probml.github.io}}$

APPLICATIONS TO PHYSICS

CARLEO ET AL.
RMP 91, 045002 (2019)

LECTURE NOTES

DAWID ET AL.
ARXIV: 2204.04198
(ESP. CHAPTER 5)

SOFTWARE

MORE IN THE TUTORIALS
(JAX, FLAX, NETKET...)

MAIN GOALS

1. KNOW ABOUT VARIATIONAL METHODS IN QUANTUM PHYSICS

2. ML - IDEAS

   a FINDING GROUND STATES

   b STUDY DYNAMICS

3. MORE "ADVANCED" TOPICS

THE MAIN PROBLEM

CASE OF A SINGLE SPIN 1/2

$$|\psi\rangle_1 = c_\uparrow |\uparrow\rangle + c_\downarrow |\downarrow\rangle$$

$$|c_\uparrow|^2 + |c_\downarrow|^2 = 1$$

CASE OF MANY SPINS

$$|\psi\rangle_N = c_{\uparrow\uparrow\ldots\uparrow} |\uparrow\uparrow\uparrow\ldots\uparrow\rangle$$

$$+ c_{\uparrow\uparrow\uparrow\ldots\downarrow} |\uparrow\uparrow\uparrow\ldots\downarrow\rangle +$$

$$+ \ldots c_{\downarrow\downarrow\downarrow\downarrow\ldots\downarrow} |\downarrow\downarrow\downarrow\ldots\downarrow\rangle$$

$$= \sum_{k=1}^{2^N} c_k |k\rangle$$

$$|k\rangle = |\sigma_1^z \ldots \sigma_N^z\rangle = |\sigma_1^z\rangle \otimes |\sigma_2^z\rangle \ldots$$

$$k \in [1, 2^N]$$

THE QUANTUM MANY-BODY PROBLEM

$\hat{H}$ SOME HAMILTONIAN

$$\hat{H}|\psi_0\rangle = E_0|\psi_0\rangle \qquad E_0 \leq E_1 \leq \dots$$

ON A COMPUTER

$$H_{ij} = \langle i|\hat{H}|j\rangle \rightsquigarrow 2^N \times 2^N$$

MATRIX

DIAGONALIZE

$|\psi_k\rangle, E_k \rightsquigarrow$ COST IS $O(2^N)$

LOCAL HAMILTONIANS

[QUANTUM INFORMATION]

$$\langle k | \hat{H} | k' \rangle \neq 0$$

↳ FIX k

\# k' IS POLY(N)

$$\hat{H} \doteq \begin{array}{cc} \begin{pmatrix} \times & \cdots & \times & \cdots & \times \end{pmatrix} & k \end{array}$$

SPARSE HAMILTONIAN MATRICES

EXAMPLE

$$\hat{H} = \sum_{i<j} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z + $$

$$+ \Gamma \sum_i^N \hat{\sigma}_i^x$$

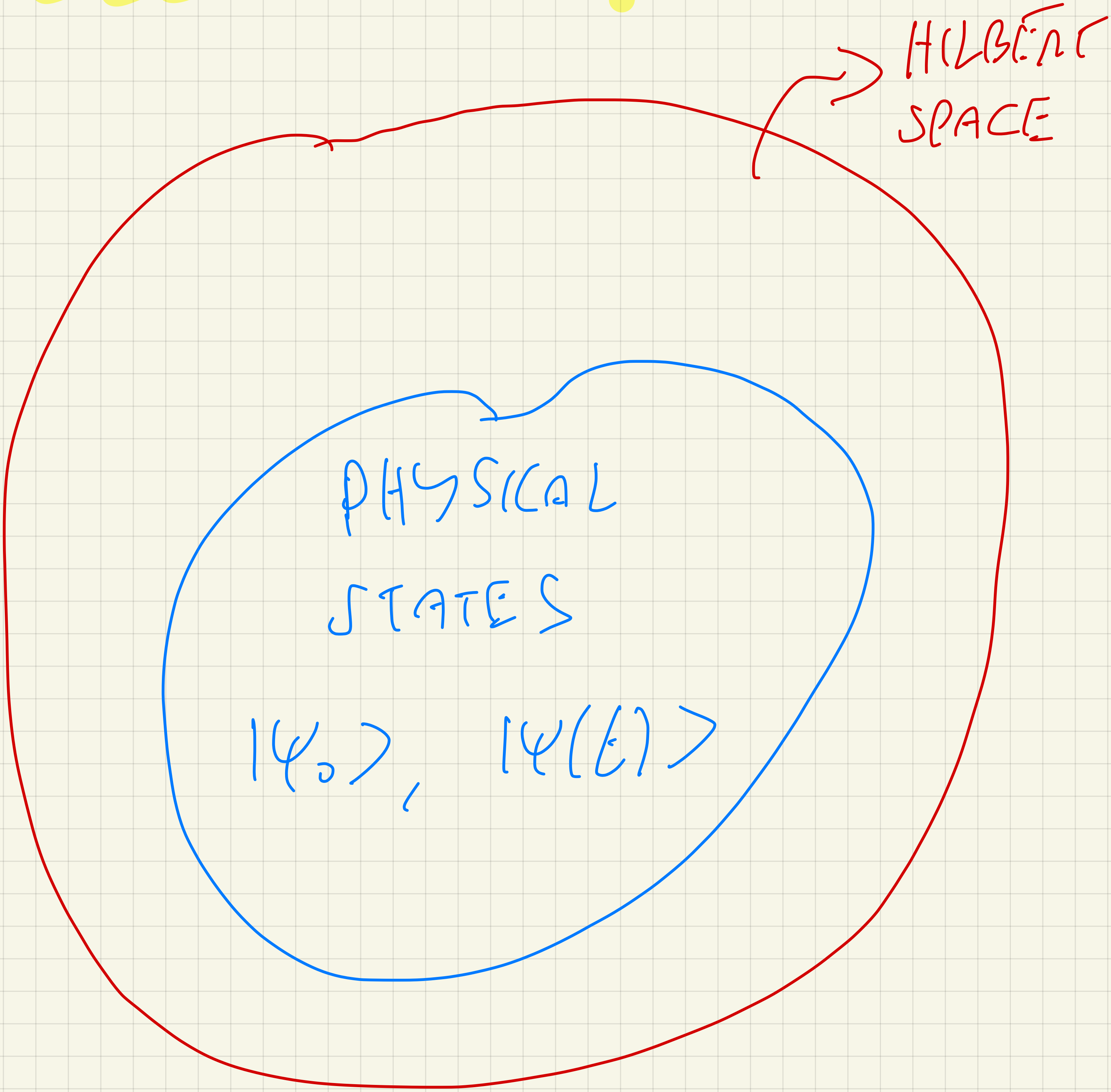
$$\langle k | \hat{\sigma}_i^z \hat{\sigma}_j^z | k \rangle \neq 0 \quad |k'\rangle = |k\rangle$$

$$\langle k | \hat{\sigma}_i^x | k' \rangle \neq 0 \quad \text{ONLY IF}$$

FOR FIXED k $\longrightarrow$ $|k'\rangle = |\sigma_1^z \dots -\sigma_i^z \dots \sigma_N\rangle$

$\#$ NON-ZERO $= (N+1)$

VARIATIONAL METHODS



ACCESSING CORNERS

$$|\psi\rangle_N = \sum_{k=1}^{2^N} C_k(\vec{\theta})\,|k\rangle$$

PARAMETERS
[ONLY POLY(N)]

FIND "OPTIMAL" PARAMETERS

$$E(\vec{\theta}) = \frac{\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle}{\langle\psi(\vec{\theta})|\psi(\vec{\theta})\rangle} \geq E_0$$

$$E(\vec{\theta}) = \frac{\sum_k |b_k|^2 E_k}{\sum_k |b_k|^2} \geq E_0$$

$$b_k = \langle\psi_k|\psi(\theta)\rangle$$

TWO MAIN FAMILIES OF STATES

1. COMPUTE $E(\vec{\theta})$ "EXACTLY" [APART FROM ROUNDING ERRORS]

2. COMPUTE $E(\vec{\theta})$ "APPROXIMATELY" [IN A CONTROLLED WAY]

EXAMPLE OF 1ST KIND

a) MEAN-FIELD STATES

$$|\psi(\vec{\theta})\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \dots |\phi_N\rangle$$

$$\langle \phi_i | \phi_j \rangle = \delta_{ij}$$

$$\langle \sigma_i^z | \phi_i \rangle = \begin{cases} \theta^{(i)}_{\uparrow} \\ \theta^{(i)}_{\downarrow} \end{cases}$$

TOTAL OF 2N PARAMETERS

e.g. $\langle \psi(\vec{\theta}) | \hat{\sigma}_i^x | \psi(\vec{\theta}) \rangle = \langle \phi_i | \hat{\sigma}_i^x | \phi_i \rangle$

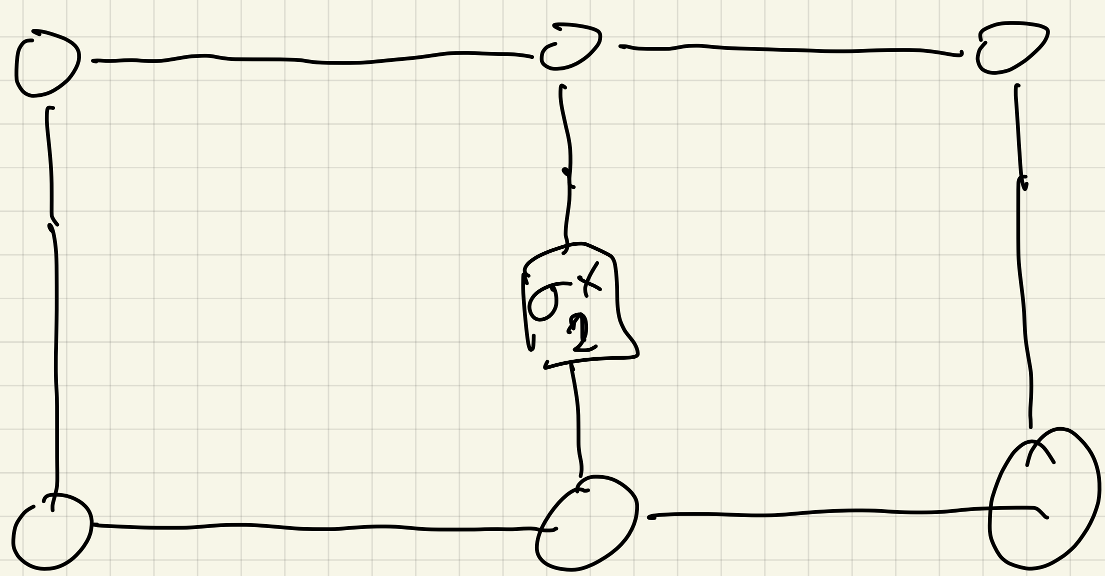
(b) MATRIX PRODUCT STATES

$$C_k(\vec{\theta}) = \hat{M}_1(\sigma_1^z) \cdot \hat{M}_2(\sigma_2^z) \cdots \hat{M}_N(\sigma_N^z)$$

EACH MATRIX $\sim D \times D$

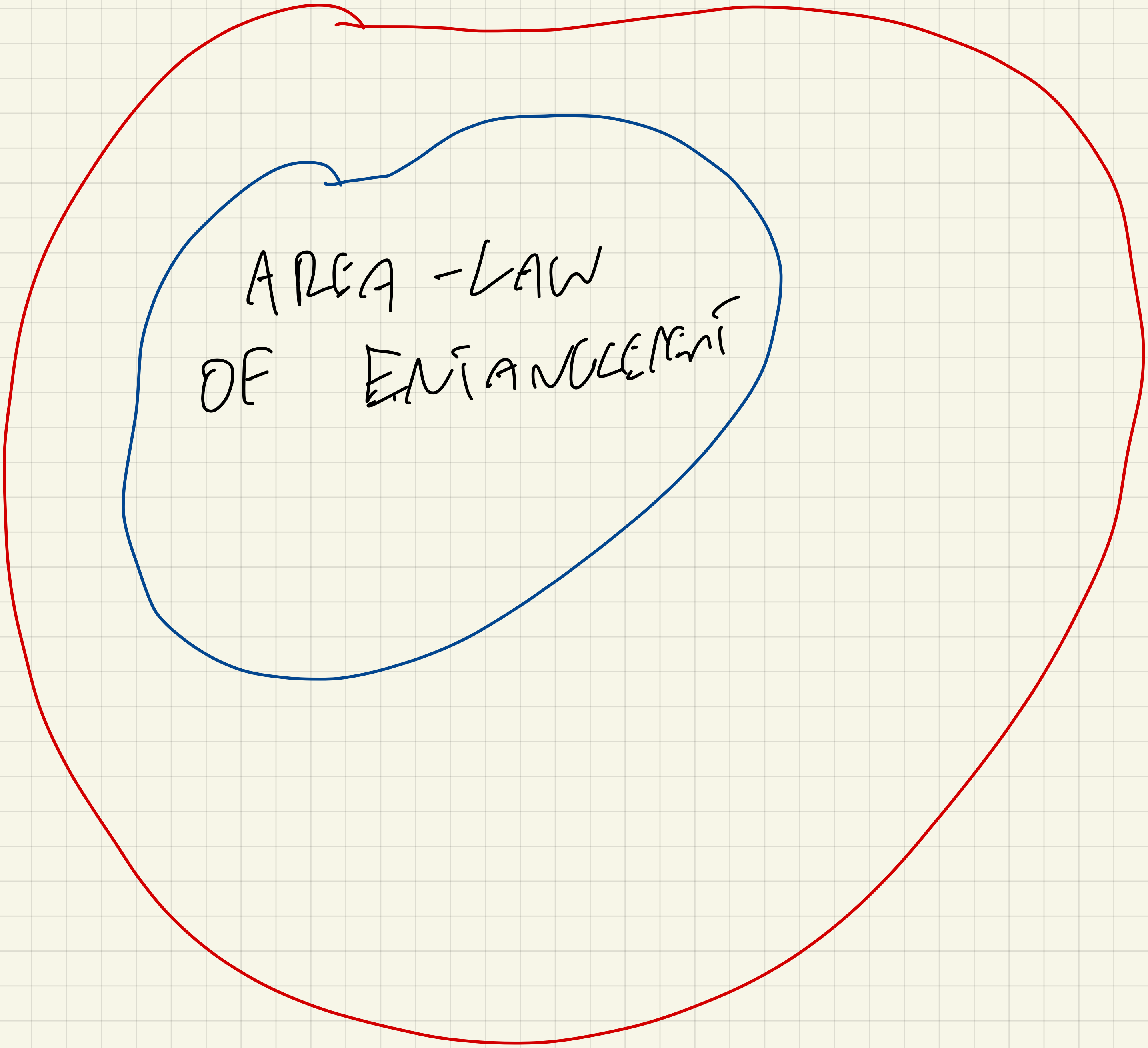
BOND DIMENSION

PARAMETERS $\simeq O(D^2 \times N)$



COMPUTING $E(\tilde{\theta}) \sim D^3$

STEMS FROM LOCALITY



SECOND TYPE OF STATES

FORMALIZED

VAN DEN NEST

ARXIV: 0911.1624

(2009)

USED SINCE

McMILLAN

~ 1960

[VARIATIONAL MONTE CARLO]

DEFINITION

"COMP. TRACTABLE STATES"

a) COMPUTE $\langle x | \psi(\vec{\theta}) \rangle$ EFFICIENTLY

b) SAMPLE FROM $P(x) = \frac{|\langle x | \psi(\vec{\theta}) \rangle|^2}{\langle \psi | \psi \rangle}$

THEOREM

$$\frac{\langle \psi(\vec{\theta}) | \hat{O} | \psi(\vec{\theta}) \rangle}{\langle \psi(\vec{\theta}) | \psi(\vec{\theta}) \rangle} \sim$$

CAN BE COMPUTED IN POLY(N) WITH FIXED ACCURACY

ESTIMATING PROPERTIES

$$\langle \hat{O} \rangle = \frac{\langle \psi | \hat{O} | \psi \rangle}{\langle \psi | \psi \rangle} =$$

$$= \frac{\sum_{k,k'} \langle \psi | k \rangle \langle k | \hat{O} | k' \rangle \langle k' | \psi \rangle}{\sum_{k} |\langle \psi | k \rangle|^2} =$$

$$= \frac{\sum_{k} |\langle \psi | k \rangle|^2 \left(\sum_{k'} \langle k | \hat{O} | k' \rangle \frac{\langle k' | \psi \rangle}{\langle k | \psi \rangle} \right)}{\sum_{k} |\langle \psi | k \rangle|^2}$$

$$= \sum_{k} P(k)\, O^{LOC}(k)$$

$$P(k) = \frac{|\langle \psi | k \rangle|^2}{\langle \psi | \psi \rangle}$$

$$O^{loc}(k) = \sum_{k'} \langle k | \hat{O} | k' \rangle \frac{\langle k' | \psi \rangle}{\langle k | \psi \rangle}$$

$$\langle \hat{O} \rangle = \mathop{E}_{k \sim P} \left[O^{loc}(k) \right]$$

STOCHASTIC ESTIMATE

① GENERATE

$$K^{(1)},\ K^{(2)},\ \dots\ K^{(M)} \sim P(K)$$

② ESTIMATE

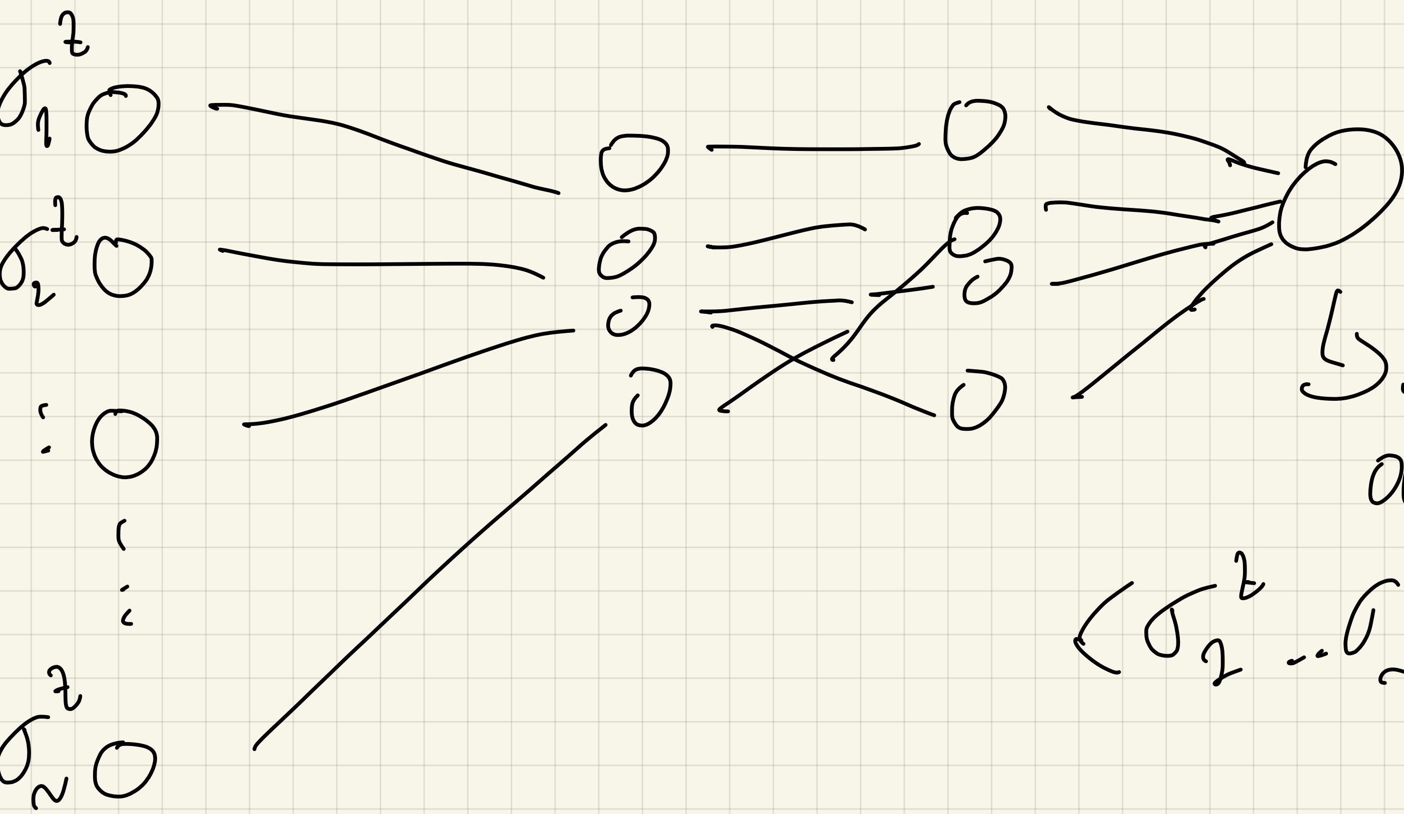
$$\langle \hat{O} \rangle \sim \frac{1}{M} \sum_{i=1}^{M} O^{LOC}\left(K^{(i)}\right)$$

ACCURACY, ERROR $\left(\langle \hat{O} \rangle_M\right) \sim \sqrt{\frac{VAR(O^{LOC})}{M}}$

REPRESENTATION OF STATE AS ANN

$$\langle k | \psi(\vec{\theta}) \rangle = \psi(\sigma_1^z, \dots \sigma_N^z; \vec{\theta})$$

$$\psi(\sigma_1^z \dots \sigma_N^z; \vec{\theta}) = g^{(D)} \circ W^{(D)} \cdots g^{(2)} \circ W^{(2)} g^{(1)} \cdot W^{(1)} \vec{\sigma}$$



↳ SINGLE OUTPUT

$$\langle \sigma_1^z \dots \sigma_N^z | \psi \rangle$$

CARLEO AND TROYER [SCIENCE 355, 602 (2017)]

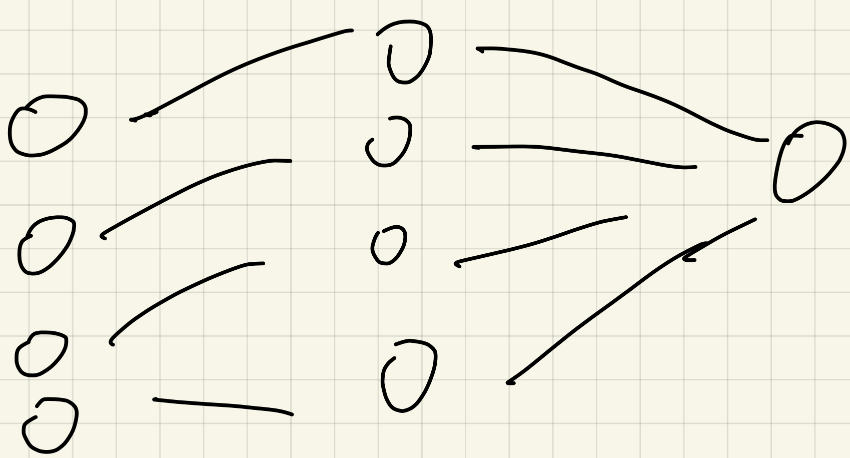
COMPLEX-VALUED OUTPUTS

1. ADOPT COMPLEX-VALUED WEIGHTS $W_{ij} \in \mathbb{C}$

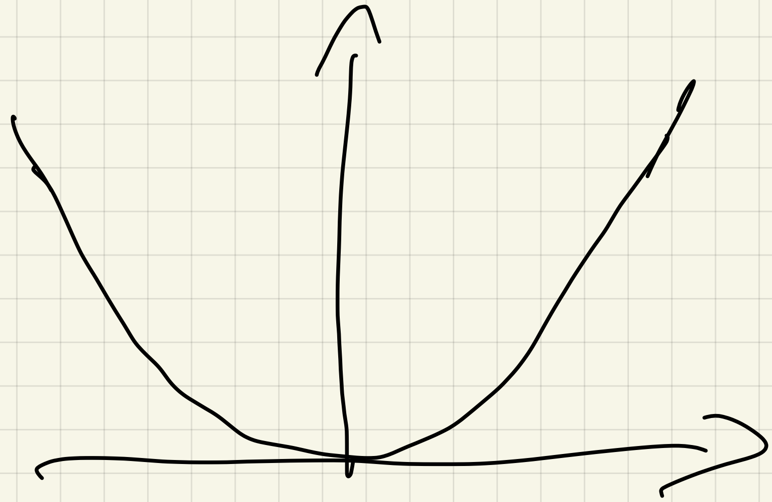
$$g(x) = \log\cosh(x)$$

$$\log P(x) = \log \prod_{j=1}^{N_h} \cosh\left(\sum_{i=1}^{N} x_i W_{ij}\right)$$

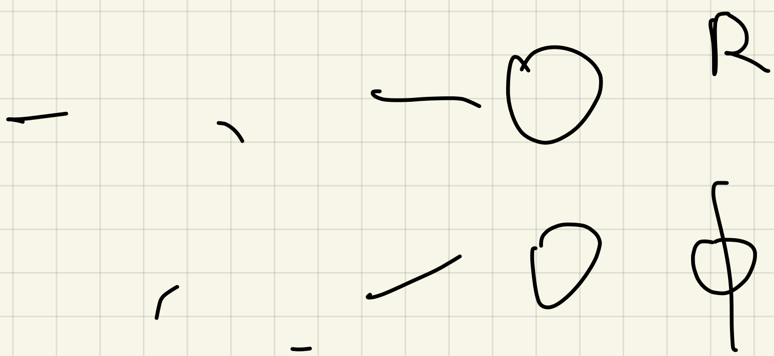
$$= \sum_{j=1}^{N_h} \log\cosh\left(\sum_{i=1}^{N} W_{ij} x_i\right)$$



$\log\cosh(x) =$



② $\langle k|\psi\rangle = e^{R(k;\theta)} \; e^{i\phi(k;\theta)}$



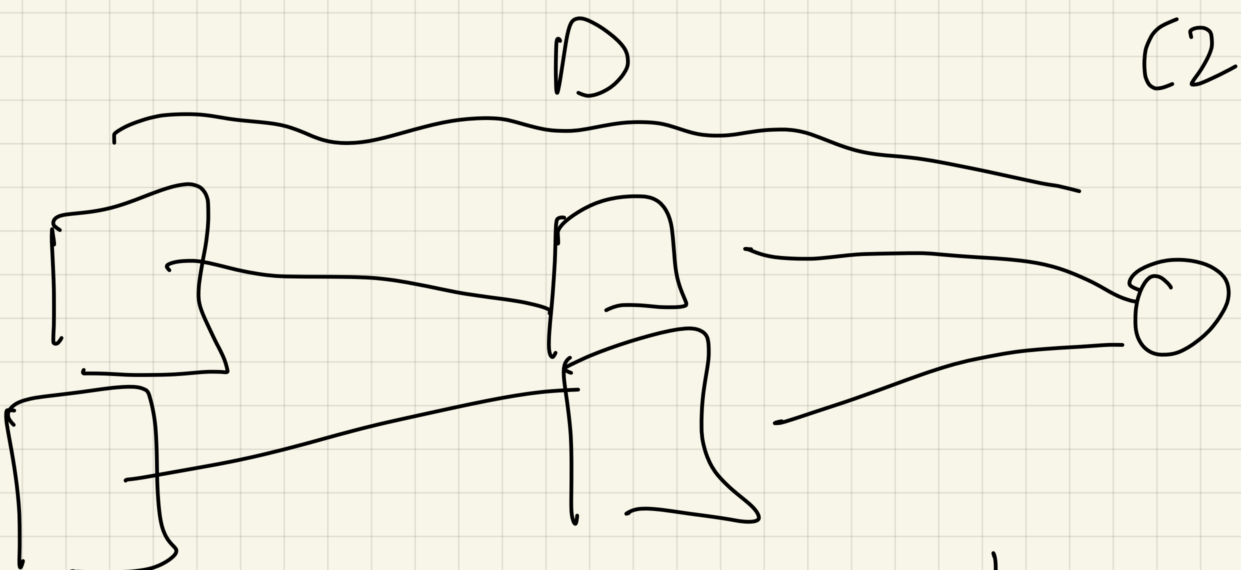
GENERAL REPR. THEOREMS

CYBENKO (1989)

KOLMOGOROV AND ARNOLD (1956)

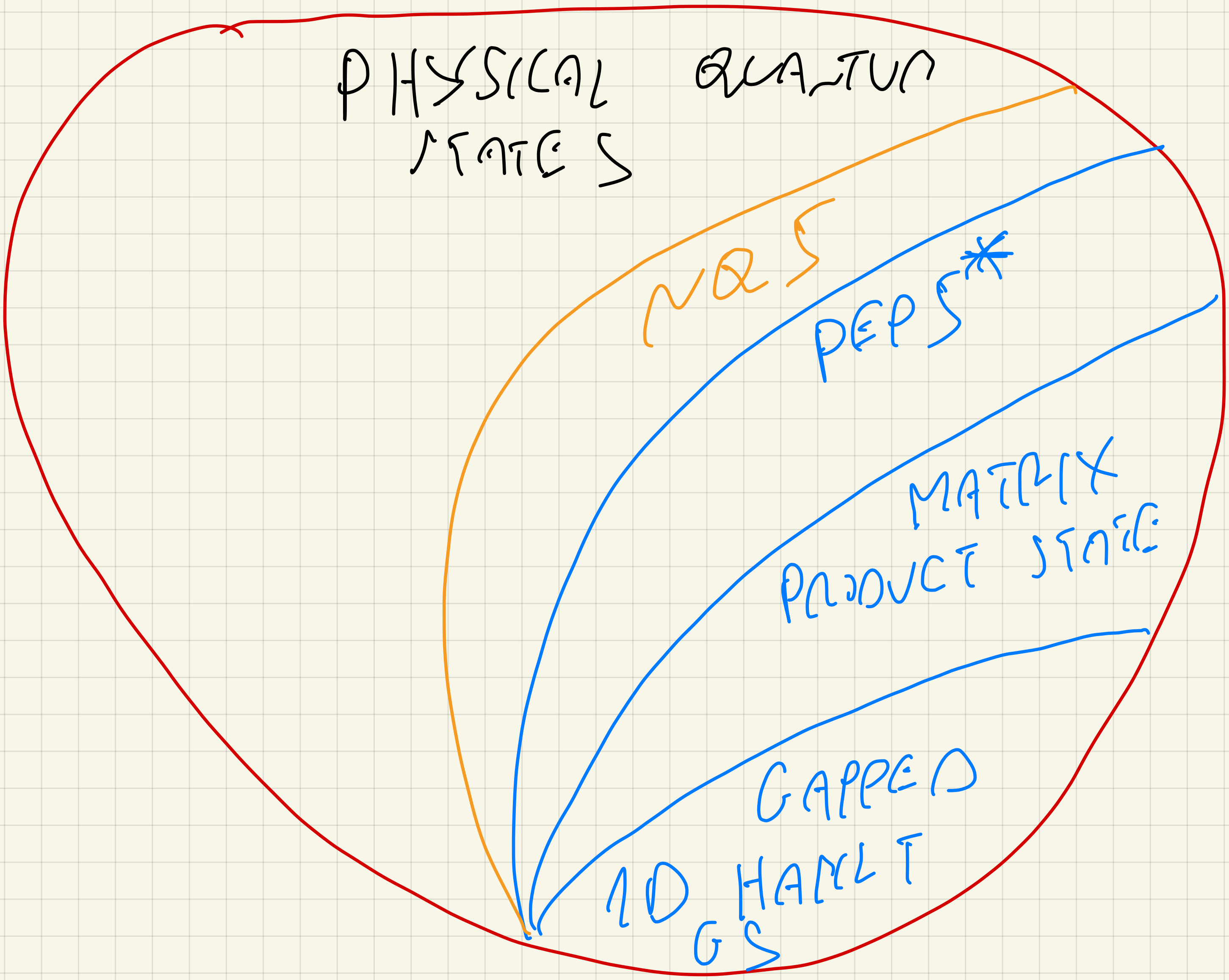
VOLUME LAW ENTANGLEMENT

LEVINE ET AL.
PRL 122, 065301 (2019)



IF $D \sim N^2$

GENERAL TENSOR NETWORKS



SHAMIR ET AL

ARXIV: 2103.10238

(2021)

LEARNING APPROACH

$$\mathcal{L}(\vec{\theta}) = E(\vec{\theta}) = \frac{\langle \psi(\vec{\theta}) | \hat{H} | \psi(\vec{\theta}) \rangle}{\langle \psi(\vec{\theta}) | \psi(\vec{\theta}) \rangle}$$

$$E(\vec{\theta}) = E_{p(k;\theta)} \left[E^{loc}(k) \right]$$

$$E^{loc}(k) = \sum_{k'} \langle k | \hat{H} | k' \rangle \frac{\langle k' | \psi \rangle}{\langle k | \psi \rangle}$$

GRADIENT CAN BE COMPUTED TOO

$$\frac{\partial E(\vec{\theta})}{\partial \theta_\ell} =$$

$$= \mathbb{E}_{P(k;\theta)}\left[E^{Loc}(k;\theta)\, O_\ell^*(k)\right]$$

$$- E(\vec{\theta})\, \mathbb{E}_{P(k;\theta)}\left[O_\ell^*(k)\right]$$

$$O_\ell(k) = \frac{\partial}{\partial \theta_\ell} \log \langle k | \psi \rangle$$

OPTIMIZATION ALGORITHM

0 INITIALIZE PARAMETERS $\vec{\theta}$

1 SAMPLE $P(k;\vec{\theta}^{(s)}) \sim k^{(1)} \dots k^{(M)}$

2 $E(\vec{\theta}^{(s)}) \simeq \frac{1}{M}\sum_i E^{loc}(k^{(i)},\vec{\theta}^{(s)})$

$G_e(\vec{\theta}^{(s)}) \simeq \langle \quad \rangle - \langle \ \rangle \langle \ \rangle$

3 $\theta_e^{(s+1)} = \theta_e^{(s)} - \eta\, G_e(\vec{\theta})$

ZERO - VARIANCE PRINCIPLE

$$E^{loc}(k) = \sum_{k'} \langle k | \hat{H} | k' \rangle \frac{\langle k' | \psi \rangle}{\langle k | \psi \rangle} =$$

$$= \frac{\langle k | \hat{H} | \psi \rangle}{\langle k | \psi \rangle}$$

if $|\psi\rangle = |\psi_0\rangle$

$$E^{loc}(k) = E_0 \frac{\cancel{\langle k | \psi \rangle}}{\cancel{\langle k | \psi \rangle}}$$

$$\operatorname*{Var}_{P(k)}\left(E^{loc}(k)\right) = \langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2$$

CAN WE SIMULATE DYNAMICS?

$$E(\vec{\theta}) = \frac{\langle \psi(\vec{\theta}) | \hat{H} | \psi(\vec{\theta}) \rangle}{\langle \psi(\vec{\theta}) | \psi(\vec{\theta}) \rangle} \geq E_0$$

TIME-DEPENDENT PROBLEMS

$$i \frac{\partial}{\partial t} | \phi \rangle = \hat{H} | \phi \rangle$$

$$| \phi(t) \rangle =$$ VARIATIONAL APPROX. ?

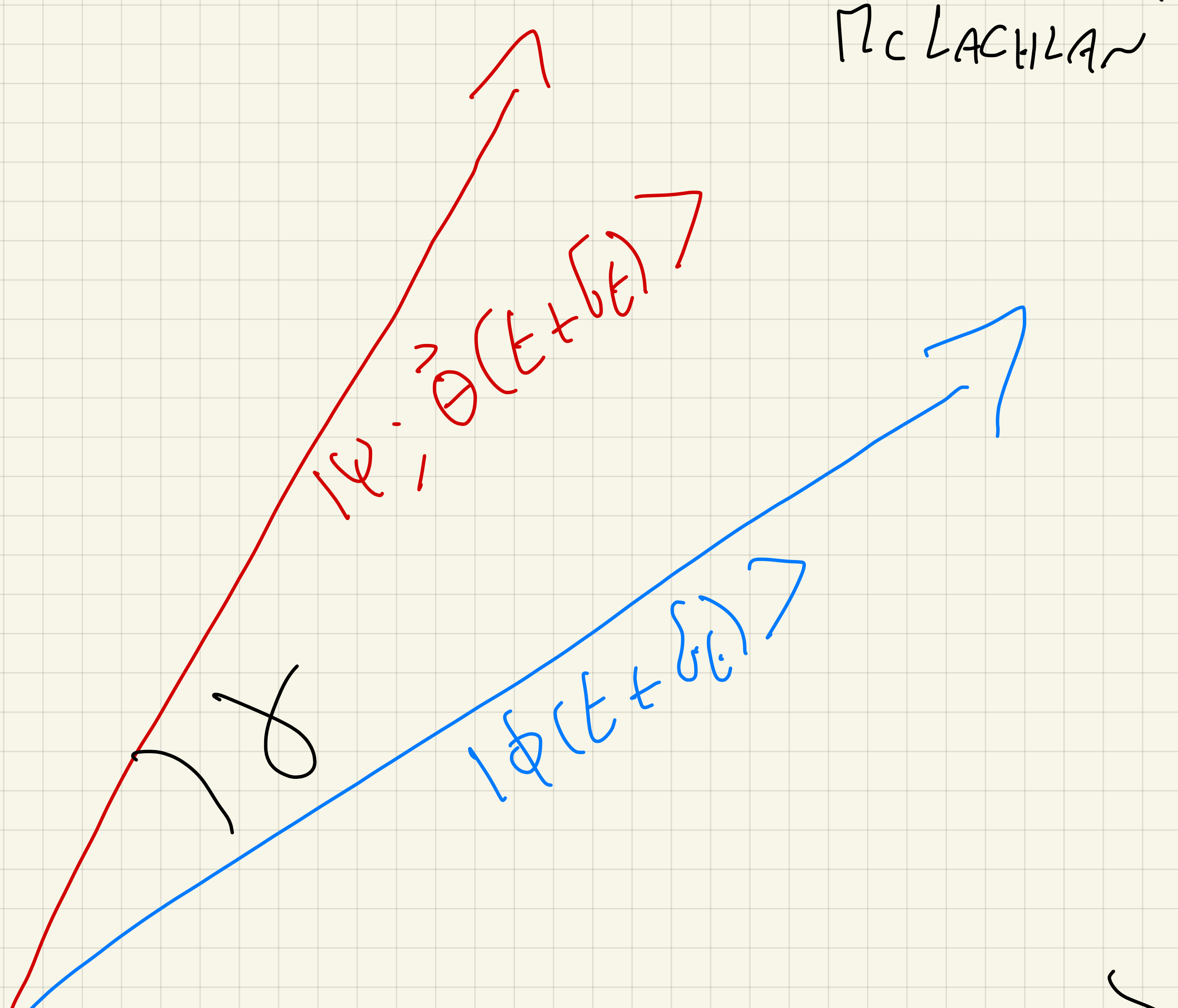
TIME-DEPENDENT VARIATIONAL ANSATZ

$$|\phi(t)\rangle \simeq |\psi;\vec{\theta}(t)\rangle$$

↳ MOVED THE TIME DEP. OF STATE INTO TIME DEP. OF PARAMETERS

TIME-DEPENDENT VARIATIONAL PRINCIPLE

[DIRAC, FRENKEL, McLACHLAN]



$$|\phi(t)\rangle = |\psi;\theta(t)\rangle$$

I $\quad |\phi(t)\rangle = |\psi;\vec{\theta}(t)\rangle$

II $\quad |\phi(t+\delta_t)\rangle = \left[\hat{1} - i\delta_t \hat{H}\right]|\phi(t)\rangle + \mathcal{O}(\delta_t^2)$

$$|\psi;\vec{\theta}(t+\delta_t)\rangle = \left[\hat{1} + \delta_t \sum_{\ell=1}^{N_p} \dot{\theta}_\ell \hat{O}_\ell\right]|\phi(t)\rangle$$

$$O_\ell(k) = \langle k|\hat{O}_\ell|k\rangle \equiv \frac{\partial}{\partial\theta_\ell}\log\psi(k)$$

$$= \frac{1}{\psi(k)}\frac{\partial}{\partial\theta_\ell}\psi(k)$$

$$F(\dot{\vec{\theta}}) = \frac{|\langle \quad | \quad \rangle|^2}{\langle \quad | \quad \rangle \langle \quad | \quad \rangle}$$

$$\max_{\dot{\theta}} F(\dot{\theta})$$

F_ℓ

//

$$i \sum_{\ell'} \underbrace{S_{\ell\ell'}}_{\text{QUANTUM GEOMETRIC TENSOR}} \dot{\theta}_{\ell'} = \langle O_\ell^* E_{loc} \rangle - \langle O_\ell^* \rangle \langle E_{loc} \rangle$$

$$S_{\ell\ell'} \simeq \left\langle \frac{\partial \psi}{\partial \theta_\ell} \middle| \frac{\partial \psi}{\partial \theta_{\ell'}} \right\rangle$$

$$S_{\ell\ell'} = \mathop{E}_{p(k)}\left[O_\ell^*(k)\, O_{\ell'}(k) \right] +$$

$$- \mathop{E}_{p}\left[O_\ell^*(k) \right] \mathop{E}_{p}\left[O_{\ell'}(k) \right]$$

$$p(k; \vec{\theta}(t)) = \frac{|\langle k | \psi(\vec{\theta}(t)) \rangle|^2}{\langle \, | \, \rangle}$$

$$\langle \partial_{\theta_\ell} \psi | \partial_{\theta_{\ell'}} \psi \rangle =$$

$$= \sum_k \langle \partial_{\theta_\ell} \psi | k \rangle \langle k | \partial_{\theta_{\ell'}} \psi \rangle =$$

$$= \sum_k \langle \psi | k \rangle O_\ell^*(k) \langle k | \psi \rangle O_{\ell'}(k) =$$

$$= \sum_k |\langle k | \psi \rangle|^2 O_\ell^*(k)\, O_{\ell'}(k)$$

TIME-DEPENDENT VARIATIONAL MONTE CARLO (t-VMC)

[SCIENTIFIC REPORTS 2, 243 (2012)]

0. $|\Psi; \vec{\theta}(0)\rangle$

I. SAMPLE

$P(k; \vec{\theta}(t))$

II. "MEASURE" $S_{ll'}, F_l$

III. SOLVE LINEAR SYSTEM $\hat{S}\dot{\vec{\theta}} = \vec{F}$

IV. $\theta_l(t+\delta_t) = \theta_l(t) + \delta_t \dot{\theta}_l$

ALSO IN IMAGINARY TIME

$$|\phi(\tau)\rangle = e^{-\tau \hat{H}} |\phi(0)\rangle$$

$$\lim_{\tau \to \infty} |\phi(t)\rangle = |\Psi_0\rangle,$$

if $|\langle \Psi_0 | \phi(0)\rangle| \neq 0$

$$\frac{\partial}{\partial \tau} |\phi(\tau)\rangle = -\hat{H} |\phi(\tau)\rangle$$

$$\sum_{\ell'} S_{\ell\ell'} \dot{\theta}_{\ell'} = -F_\ell$$

EQUIVALENT TO "STOCHASTIC RECONFIGURATION"

[SORELLA, BECCA "QUANTUM MONTE CARLO"]

How EXPENSIVE ?

$$O(N_p^2)$$

$$O(N_p)$$

CONJUGATE GRADIENT-LIKE SOLVERS

$$\tilde{S}_{\ell\ell'} = S_{\ell\ell'} + \delta_{\ell\ell'}\Lambda$$

$$\Lambda \sim 10^{-4}$$