Reducing the Circuit Footprint for Noise-Resilient Quantum Chemistry via Transcorrelation and Adaptive Ansätze

Werner Dobrautz

Quantinuum Seminar

January 31, 2025

INSTITUTE OF

PARTICIPATING INSTITUTIONS

FUNDED BY

Federal Ministry of Education and Research

Diese Maßnahme wird mitfinanziert mit Steuermitteln auf Grundlage des vom Sächsischen Landtag beschlossenen Haushalts.

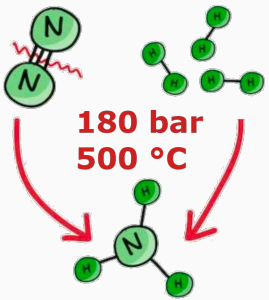
Outline

- Background and Motivation
- Transcorrelation to reduce the computational footprint on quantum hardware
- Results: Hubbard model and *ab initio* quantum chemistry problems
  - Hubbard model – Reduce circuit depth
  - *Ab initio* problems – Reduce circuit width (qubits)
- Adaptive Circuit Ansätze
- Conclusions

Background and Motivation

Motivation: Haber-Bosch process and biological nitrogen fixation

Haber-Bosch Process



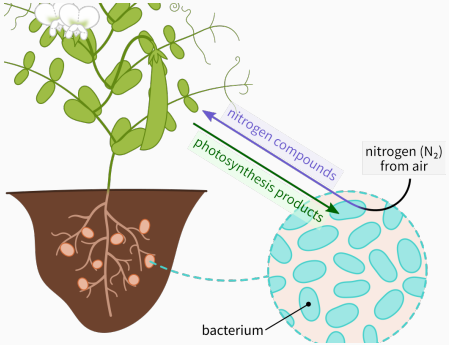
- **Crucial for fertilizer production**
- 2% of world's energy consumption
- 3% of global carbon emissions
- 5% of natural gas consumption

Motivation: Haber-Bosch process and biological nitrogen fixation

Haber-Bosch Process

- **Crucial for fertilizer production**
- 2% of world's energy consumption
- 3% of global carbon emissions
- 5% of natural gas consumption

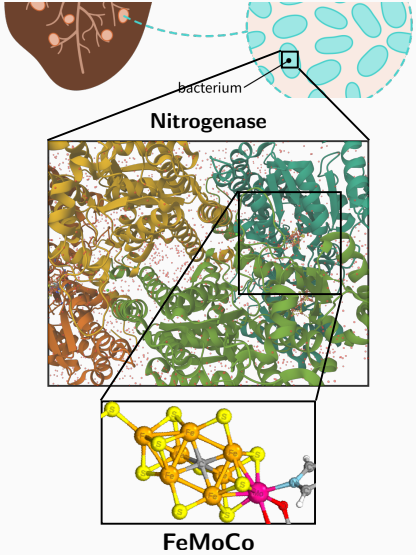
Biological nitrogen fixation



- **Ambient pressure and temperature**
- Not yet understood → Bio-catalysts for more efficient and greener ammonia production

Problem: Strongly correlated transition metal compounds

- **Transition metal clusters** act as catalysts: Iron-Molybdenum cofactor (FeMoCo) and other iron-sulfur clusters
- **Experimental study very difficult!**

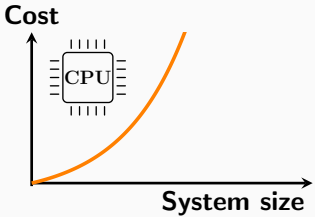


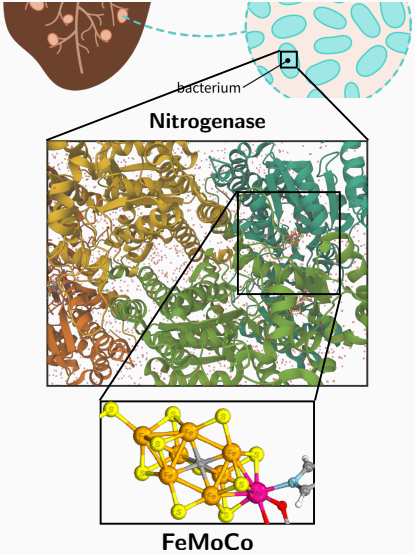
Problem: Strongly correlated transition metal compounds

- **Transition metal clusters** act as catalysts: Iron-Molybdenum cofactor (FeMoCo) and other iron-sulfur clusters
- **Experimental study very difficult!**

→ Numerical studies of relevant quantum phenomena necessary → $\hat{\mathbf{H}}\,|\mathbf{\Psi}\rangle = \mathbf{E}\,|\mathbf{\Psi}\rangle$

We have the equations at hand, but **exponentially costly** on classical computers!



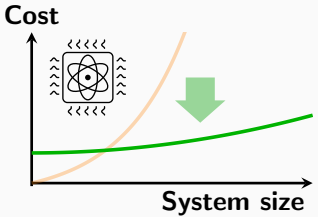


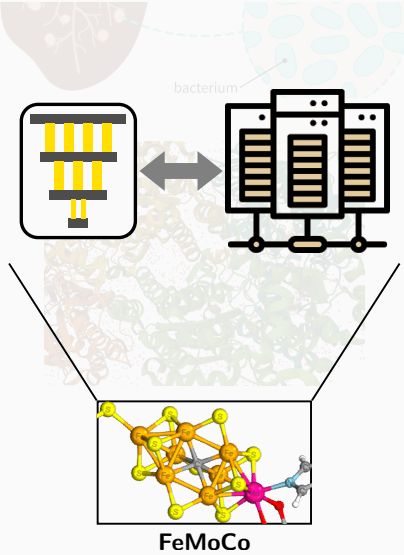
Problem: Strongly correlated transition metal compounds

- **Transition metal clusters** act as catalysts: Iron-Molybdenum cofactor (FeMoCo) and other iron-sulfur clusters
- **Experimental study very difficult!**

→ Numerical studies of relevant quantum phenomena necessary → $\hat{\mathbf{H}}\,|\Psi\rangle = \mathbf{E}\,|\Psi\rangle$

Use a combined **HPC and QC** approach for a potential computational speedup



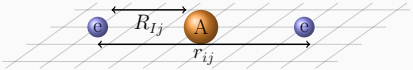


To obtain insight on the **chemical** and **physical properties** of these systems we need to **solve the Schrödinger equation**

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\,, \qquad \hat{H}\,|\Psi(t)\rangle = i\frac{\partial}{\partial t}\,|\Psi(t)\rangle$$

All necessary information contained in electronic **molecular Hamiltonian**

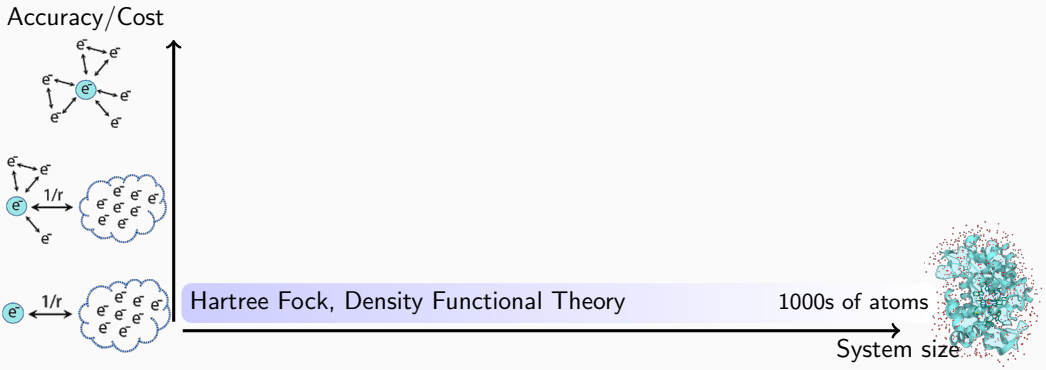
$$\hat{H} = \underbrace{-\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i}}_{\text{kinetic energy of e}^-} \underbrace{-\sum_{I,j}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_j|}}_{\text{Attr. potential}} + \underbrace{\frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\text{e}^- - \text{e}^- \text{ repulsion}}$$



Coulomb repulsion correlates all electrons of a system → analytic solution too complex → **approximations and computational approaches**

Hierarchy of Methods and Theories

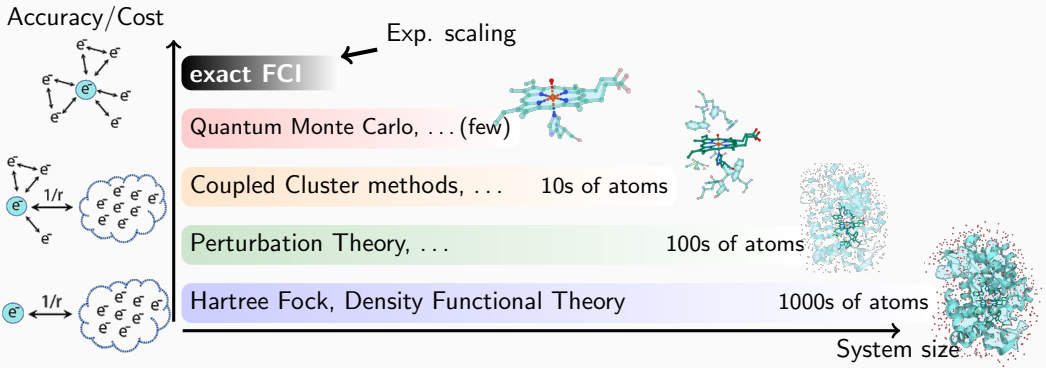
Depending how accurately we treat correlation: various methods and **levels of theory** to solve the Schrödinger equation



PRL 121, 056401 (2018); **JCP 153**, 034107 (2020); **JCTC 19**, 20, 6933 (2023)

Hierarchy of Methods and Theories

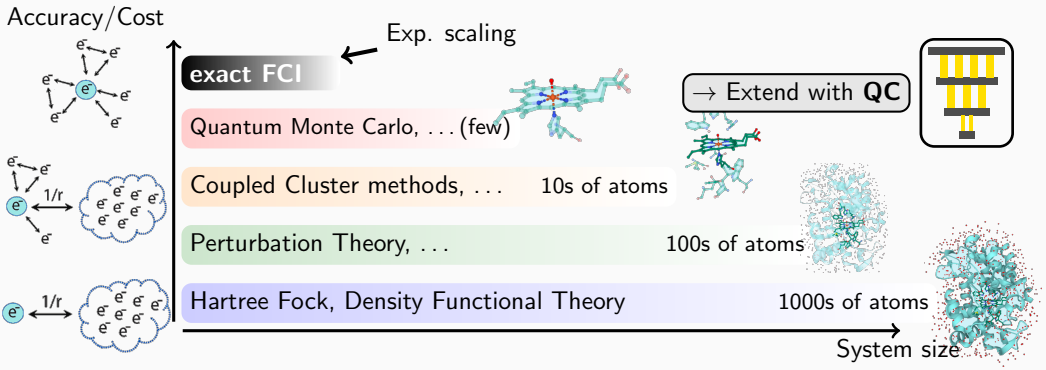
Depending how accurately we treat correlation: various methods and **levels of theory** to solve the Schrödinger equation



PRL 121, 056401 (2018); **JCP 153**, 034107 (2020); **JCTC 19**, 20, 6933 (2023)

Hierarchy of Methods and Theories

Depending how accurately we treat correlation: various methods and **levels of theory** to solve the Schrödinger equation

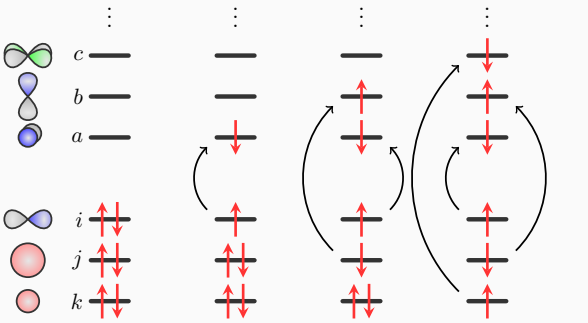


PRL 121, 056401 (2018); **JCP 153**, 034107 (2020); **JCTC 19**, 20, 6933 (2023)

Exponential scaling of Full Configuration Interaction

FCI $\Rightarrow$ exact solution in a given basis: linear combination of determinants

$$|\Psi\rangle = |\Phi_{HF}\rangle + \sum_i c_i |\Phi_i\rangle$$



All possible excitations from HF determinant

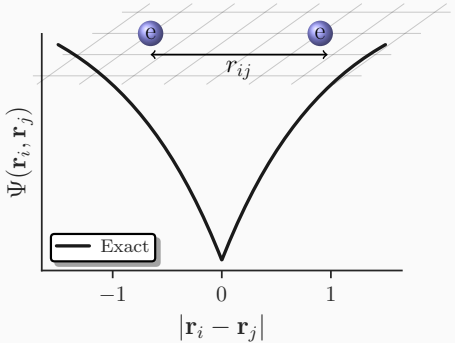
Number of possible states for given number of electrons, N, and orbitals, n, $\sim \binom{N}{n}$

Mol.	#orbitals	#electrons	#states
H_2	2	2	4
LiH	4	4	36
Be_2	8	8	4900
H_2O	12	12	$\sim 8 \cdot 10^5$
C_2H_4	16	16	$\sim 16 \cdot 10^6$
F_2	18	18	$\sim 2 \cdot 10^9$

Problems for accurate description: Cusp condition

Cusp condition: Singularity of Coulomb potential ($\sim \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$)
$\rightarrow$ sharp cusp of exact wavefunction $\Psi(\{\mathbf{r}\})$ at electron coalescence ($|\mathbf{r}_i - \mathbf{r}_j| = 0$)

$$\hat{H} \propto -\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \qquad \hat{H}\,|\Psi(\{\mathbf{r}\})\rangle = E_0\,|\Psi(\{\mathbf{r}\})\rangle$$

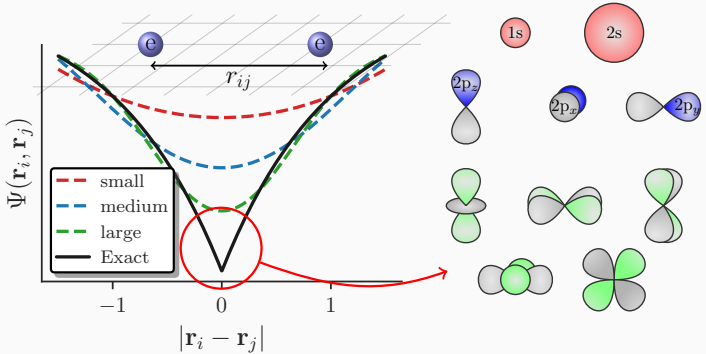


Kato, Communications on Pure and Applied Mathematics **10** (2), 151 (1957)

Problems for accurate description: Cusp condition

Cusp condition: Singularity of Coulomb potential ($\sim \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$)
$\rightarrow$ sharp cusp of exact wavefunction $\Psi(\{\mathbf{r}\})$ at electron coalescence ($|\mathbf{r}_i - \mathbf{r}_j| = 0$)

$$\hat{H} \propto -\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \qquad \hat{H}\,|\Psi(\{\mathbf{r}\})\rangle = E_0\,|\Psi(\{\mathbf{r}\})\rangle$$

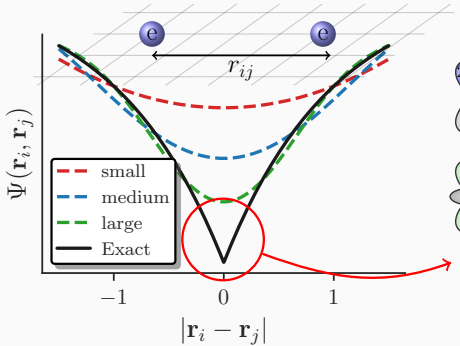


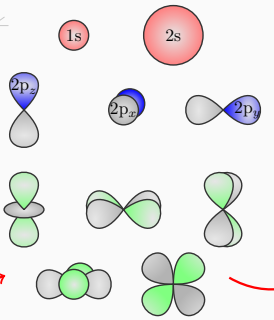
Kato, Communications on Pure and Applied Mathematics **10** (2), 151 (1957)

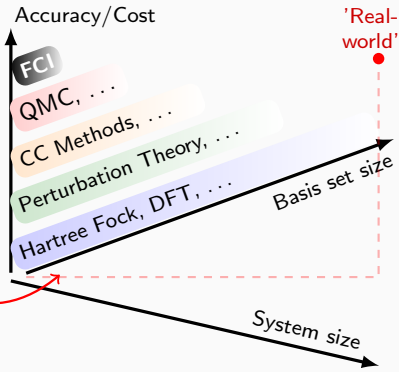
Problems for accurate description: Cusp condition

Cusp condition: Singularity of Coulomb potential ($\sim \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$)
$\rightarrow$ sharp cusp of exact wavefunction $\Psi(\{\mathbf{r}\})$ at electron coalescence ($|\mathbf{r}_i - \mathbf{r}_j| = 0$)

$$\hat{H} \propto -\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \qquad \hat{H}\,|\Psi(\{\mathbf{r}\})\rangle = E_0\,|\Psi(\{\mathbf{r}\})\rangle$$





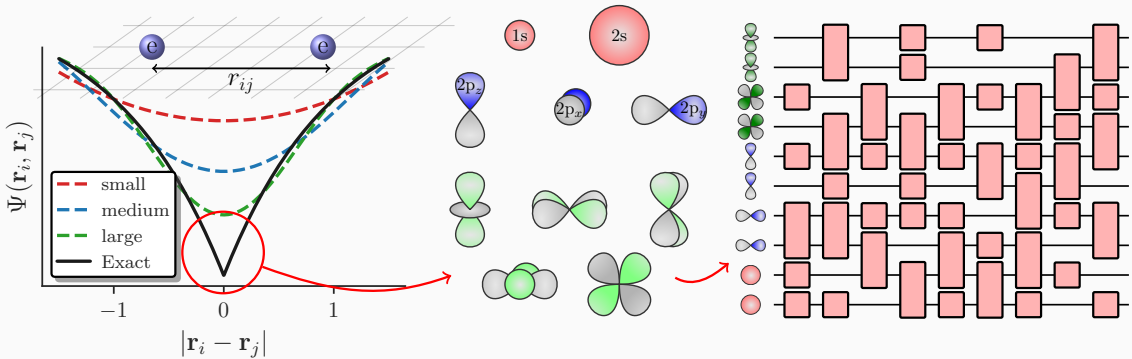


Kato, Communications on Pure and Applied Mathematics **10** (2), 151 (1957)

Problems for accurate description: Cusp condition

Cusp condition: Singularity of Coulomb potential ($\sim \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$)
$\rightarrow$ sharp cusp of exact wavefunction $\Psi(\{\mathbf{r}\})$ at electron coalescence ($|\mathbf{r}_i - \mathbf{r}_j| = 0$)

$$\hat{H} \propto -\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \qquad \hat{H}\,|\Psi(\{\mathbf{r}\})\rangle = E_0\,|\Psi(\{\mathbf{r}\})\rangle$$

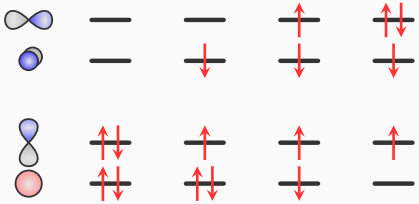


Kato, Communications on Pure and Applied Mathematics **10** (2), 151 (1957)

The Case for Quantum

How can Quantum Computing help Quantum Chemistry?

2^N states

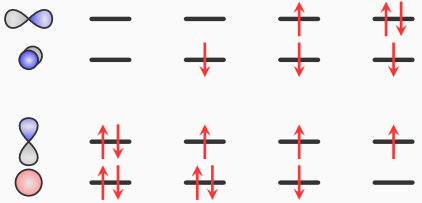


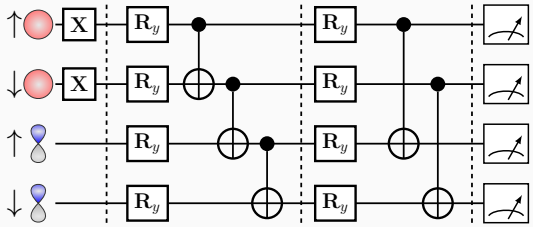
$|\Phi_{\mathrm{HF}}\rangle = |1100\rangle\,, |\Phi_i^a\rangle = |1010\rangle\,, \ldots$

How can Quantum Computing help Quantum Chemistry?

Fermion to qubit mapping

2^N states ⟶ N qubits





$|\Phi_{\mathrm{HF}}\rangle = |1100\rangle\,, |\Phi_i^a\rangle = |1010\rangle\,, \ldots$

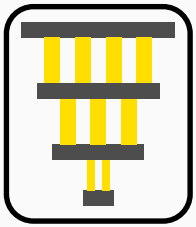
- Map our problem (Hamiltonian/basis functions) onto quantum hardware/qubits
  - Qubits encode occupation of spin-orbitals $\in [0, 1]$

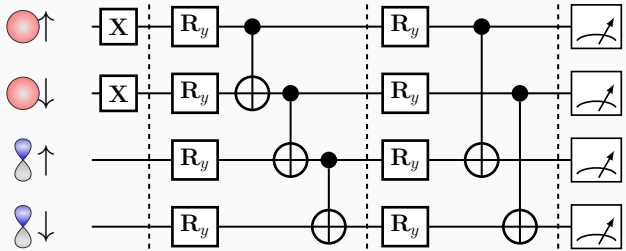
$\rightarrow$ Use quantum algorithms for ground-, excited states, dynamics, ...

* Jordan-Wigner, Bravyi-Kitaev, (Ann. Phys. **298**, 210 (2002)), Parity encoding ...

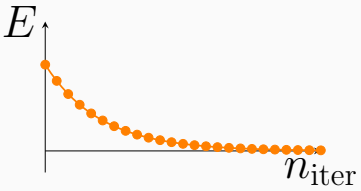
$\sim$ 15mK

Magnetic shielding

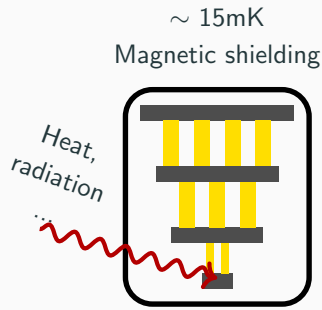


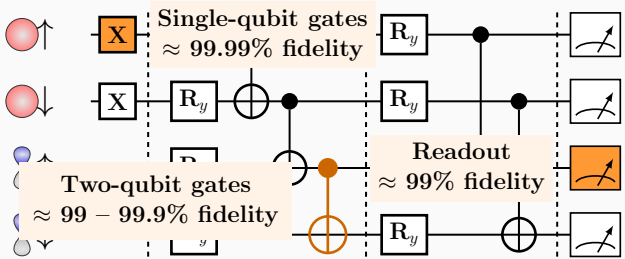


Variational Quantum Eigensolver:



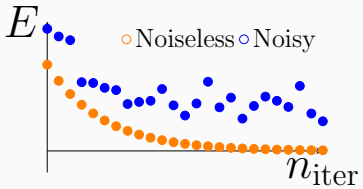
Quantum Chemistry on Quantum Computers



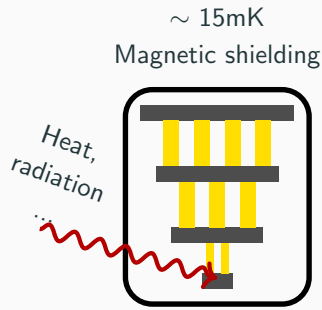


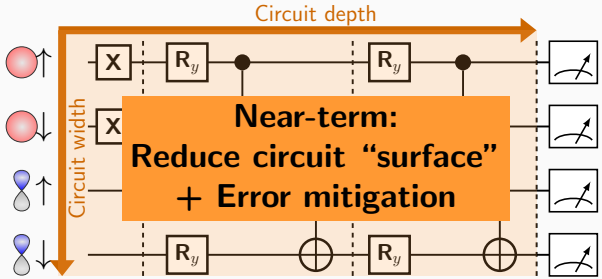
Effect of noise:

- Bit flip: $|0\rangle \leftrightarrow |1\rangle$
- Phase flip: $|0\rangle \leftrightarrow -|0\rangle$
- Decoherence: $|0\rangle + |1\rangle \rightarrow |0\rangle + e^{i?}|1\rangle$
- ...



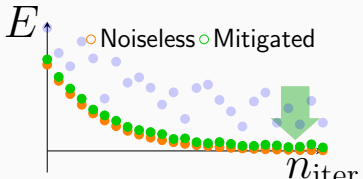
Quantum Chemistry on Quantum Computers





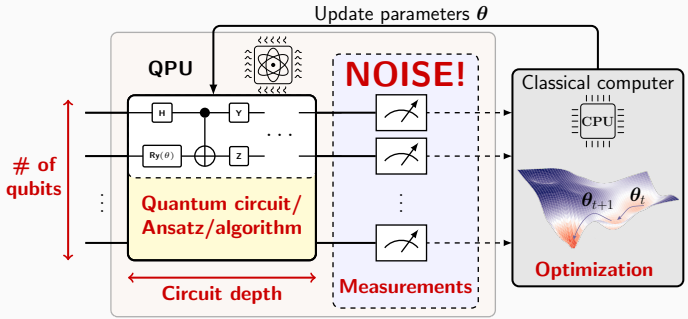
Effect of noise:

- Bit flip: $|0\rangle \leftrightarrow |1\rangle$
- Phase flip: $|0\rangle \leftrightarrow -|0\rangle$
- Decoherence: $|0\rangle + |1\rangle \rightarrow |0\rangle + \mathrm{e}^{i?}|1\rangle$
- ...



Near-term approaches and our work

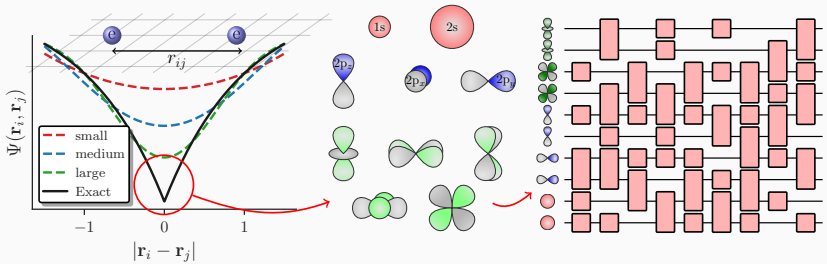
Hybrid Quantum-Classical: Use benefits of both quantum and classical resources



- Algorithms:
  Quantum imaginary time evolution (QITE)
- Classical optimization
- Resource reduction:
  Qubits and circuit depth
- Error mitigation

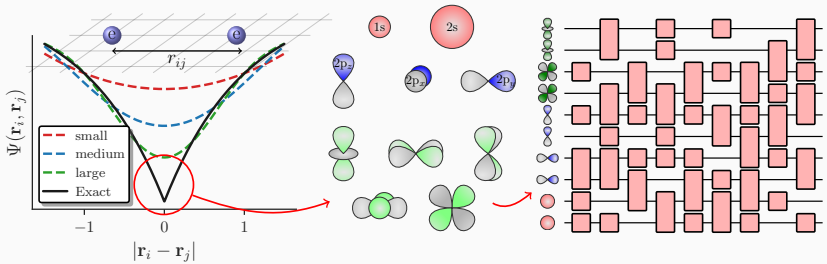
Transcorrelation to reduce the computational footprint on quantum hardware

Resource Reduction: Qubits and circuit depth

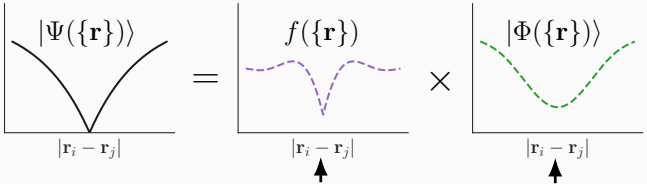


Form of the **cusp** is known* $\rightarrow$ describe it with a **wavefunction Ansatz**

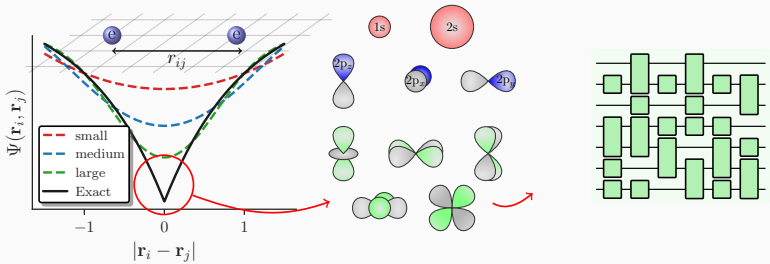
Resource Reduction: Qubits and circuit depth



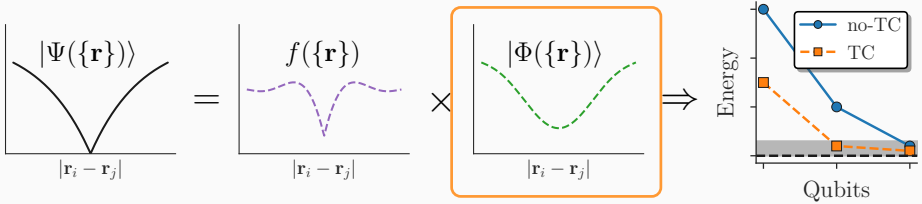
Form of the **cusp** is known* $\rightarrow$ describe it with a **wavefunction Ansatz**



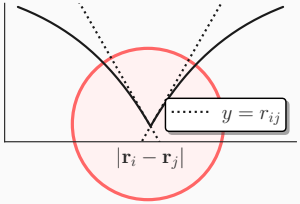
Resource Reduction: Qubits and circuit depth



Form of the **cusp** is known* $\rightarrow$ describe it with a **wavefunction Ansatz**



Explicitly Correlated methods



Linear behavior in electron-electron distance $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ for small r_{ij}!

R12 methods*: $|\Psi\rangle = r_{ij}\,|\Phi\rangle$

F12 methods†: $|\Psi\rangle = f(r_{ij})\,|\Phi\rangle\,, \quad f(r_{ij}) = \dfrac{1-\exp(-\gamma r_{ij})}{\gamma}$

Jastrow Ansatz‡: $\quad |\Psi\rangle = \mathrm{e}^{\hat{J}}\,|\Phi\rangle\,, \quad \hat{J} = \sum_{ij} J_{ij} g(\tilde{r}_{ij})$

J_{ij} are optimizable parameters and $g(\tilde{r_{ij}})$ polynomials dependent on the electron positions.
We use VMC# to classically optimize the Jastrow factor $\hat{J}$, which scales as $\mathcal{O}(n_{\mathrm{el}}^3)$

* Kutzelnigg, Theoretica chimica acta 68, 445 (1985); † Ten-no, J. Chem. Phys. 121, 117 (2004); ‡ Jastrow, Phys. Rev. **98**, 1479 (1955); #Haupt, Hosseini, López Ríos, **WD**, Cohen and Alavi, JCP **158**, 224105 (2023);

Cusp Condition – The Transcorrelated (TC) Method

The transcorrelated (TC) method: use a Jastrow Ansatz, $\mathrm{e}^{\hat{J}}$, with optimizable parameters J_{ij} (via VMC†) to describe the cusp condition and/or capture part of correlation.

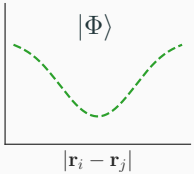
$$|\Psi(\{\mathbf{r}\})\rangle = \exp\left[\sum_{ij} J_{ij} g(\tilde{r}_{ij})\right] |\Phi(\{\mathbf{r}\})\rangle$$

Incorporate $\mathrm{e}^{\hat{J}}$ into Hamiltonian and solve a similarity transformed problem:

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\,, \quad \text{with} \quad |\Psi\rangle = \mathrm{e}^{\hat{J}}\,|\Phi\rangle$$

$$\mathrm{e}^{-\hat{J}} \rightarrow | \quad \hat{H}\,\mathrm{e}^{\hat{J}}\,|\Phi\rangle = E\,\mathrm{e}^{\hat{J}}\,|\Phi\rangle\,, \qquad \left(\hat{J}^{\dagger} = \hat{J}\right)$$

$$\left(\mathrm{e}^{-\hat{J}}\,\hat{H}\,\mathrm{e}^{\hat{J}}\right)|\Phi\rangle = E\,\mathrm{e}^{-\hat{J}}\,\mathrm{e}^{\hat{J}}\,|\Phi\rangle = E\,|\Phi\rangle$$



$|\Phi\rangle$ easier to represent with less basis functions $\rightarrow$ immense resource reduction

* Kato (1957); Boys and Handy (**1969**); Kutzelnigg (1985); **WD**, Luo, Alavi, PRB **99** (7), 075119 (2019); Cohen, Luo, Guther, **WD**, Tew, Alavi, JCP **151** (6), 061101 (2019); **WD**, Cohen, Alavi, Giner, JCP **156** (23), 234108 (2022); † Haupt, Hosseini, López Ríos, **WD**, Cohen and Alavi, JCP **158**, 224105 (2023);

Similarity Transformation – Transcorrelated (TC) Method

Baker-Campbell-Hausdorff (BCH) exp. to obtain TC Hamiltonian:

$$\bar{H} = \mathrm{e}^{-\hat{J}}\,\hat{H}\,\mathrm{e}^{\hat{J}} = \hat{H} + [\,\hat{H},\hat{J}\,] + \frac{1}{2}[\,[\,\hat{H},\hat{J}\,],\hat{J}\,] + \ldots$$

For the molecular Hamiltonian the BCH exp. terminates at 2nd order, as only kinetic energy operators in $\hat{H}$ do not commute with $\hat{J}$!

$$\begin{aligned}\bar{H} &= \hat{H} - \sum_i \left(\frac{1}{2}\nabla_i^2\hat{J} + (\nabla_i\hat{J})\nabla_i + \frac{1}{2}(\nabla_i\hat{J})^2\right) \\ &= \hat{H} - \sum_{i<j}\hat{K}(\mathbf{r}_i,\mathbf{r}_j) - \sum_{i<j<k}\hat{L}(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k)\end{aligned}$$

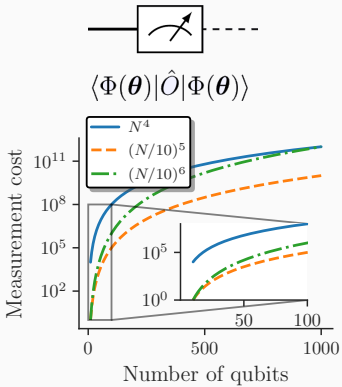
- Exact transformation
- 3-body terms and non-Hermitian!
- **Rapid basis set convergence!**

Hirschfelder, JCP, **39**, 3145 (1963); Boys and Handy, Proc. R. Soc. A (**1969**); **WD**, Luo, Alavi, PRB **99** (7), 075119 (2019); Cohen, Luo, Guther, **WD**, Tew, Alavi, JCP **151** (6), 061101 (2019); **WD**, Cohen, Alavi, Giner, JCP **156** (23), 234108 (2022)

Scaling of TC – Measurement Cost

$$\bar{H} = \sum_{pq,\sigma} h_q^p a_{p,\sigma}^\dagger a_{q,\sigma} + \frac{1}{2} \sum_{pqrs,\sigma\tau} \bar{V}_{rs}^{pq} a_{p,\sigma}^\dagger a_{q,\tau}^\dagger a_{s,\tau} a_{r,\sigma} - \frac{1}{6} \sum_{pqrstu,\sigma\tau\lambda} L_{stu}^{pqr} a_{p,\sigma}^\dagger a_{q,\tau}^\dagger a_{r,\lambda}^\dagger a_{u,\lambda} a_{t,\tau} a_{s,\sigma}$$

- Measurement formally scaling as N^6, with N being the number of orbitals
- Shown that N^6-scaling terms can be neglected to good accuracy*
- **xTC work on N^4-scaling approximation**‡
- **Order of magnitude less orbitals:** since also no core functions needed in basis set†
- **Shorter circuit depth**, due to more compact ground state!#



*WD *et al.*, Journal of Chemical Physics 156 (23), 234108 (2022); †Cohen, Luo, Guther, **WD**, Tew, Alavi, JCP **151** (6), 061101 (2019); ‡Christlmaier, Schraivogel, López Ríos, Alavi, Kats, JCP **159**, (1) 014113 (2023); #Sokolov, **WD**, Luo, Alavi, Tavernelli, PR Research **5** (2), 023174 (2023);

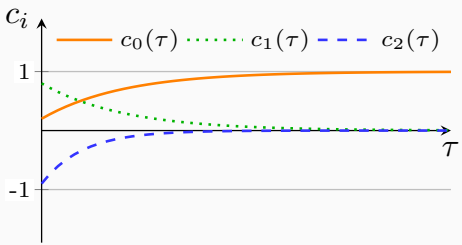
Since the TC Hamiltonian is non Hermitian, variational algorithms like VQE are not applicable!

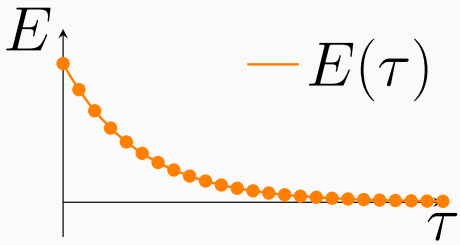
$\rightarrow$ Quantum Imaginary Time Evolution!

(Quantum) Imaginary Time Evolution – QITE

→ Solve for the **right** eigenvector of non-Hermitian $\bar{H}$ by projecting on the ground state with **(quantum) imaginary-time evolution (QITE)**

$$i\frac{\partial |\Psi\rangle}{\partial t} = \hat{H}|\Psi\rangle \quad \overset{\tau=it}{\rightarrow} \quad \frac{\partial |\Psi\rangle}{\partial \tau} = -\hat{H}|\Psi\rangle \quad \rightarrow \quad \boxed{|\Psi_0\rangle = \lim_{\tau\to\infty} \mathrm{e}^{-\hat{H}\tau}|\Phi(0)\rangle}$$





Variational Ansatz-based QITE – VarQITE

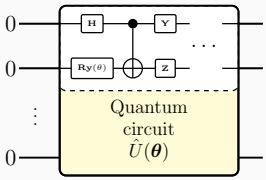
(Normalized) imaginary-time (Wick-rotated) Schrödinger equation

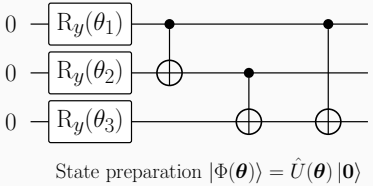
$$\frac{\partial \left|\Psi(\tau)\right\rangle}{\partial \tau} = -(\hat{H} - E_\tau)\left|\Psi(\tau)\right\rangle \quad \rightarrow \left|\Psi(\tau)\right\rangle = \mathrm{e}^{-\delta\tau(\hat{H}-E_\tau)}\left|\Psi(0)\right\rangle$$

$$\text{with} \quad E_\tau = \left\langle\Psi(\tau)\right| \hat{H} \left|\Psi(\tau)\right\rangle$$

would yield the ground state, but **non-unitary** $\mathrm{e}^{-\tau(\hat{H}-E_\tau)}$ not possible on a QC!
Approximate $|\Psi(\tau)\rangle$ with an "Ansatz" with parametrized unitary gates:

$$\left|\Psi(\tau)\right\rangle \approx \left|\Phi(\boldsymbol{\theta}(\tau))\right\rangle = \hat{U}(\boldsymbol{\theta}(\tau))\left|\mathbf{0}\right\rangle$$



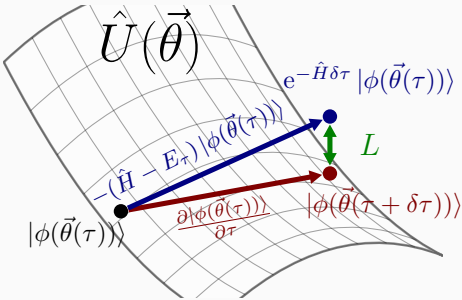


State preparation $\left|\Phi(\boldsymbol{\theta})\right\rangle = \hat{U}(\boldsymbol{\theta})\left|\mathbf{0}\right\rangle$

McArdle *et al.*, npj Quantum, 5, 75 (2019); Yuan *et al.* Quantum 3, 191 (2019);

VarQITE – Details

Map imaginary-time evolution to parameters $\boldsymbol{\theta}(\tau)$ of Ansatz $|\Phi(\boldsymbol{\theta}(\tau))\rangle \approx |\Psi(\tau)\rangle$



- Imag-time Schrödinger equation, for small $\delta\tau$:

$$e^{-\delta\tau(\hat{H}-E_\tau)} \approx (1-\delta\tau(\hat{H}-E_\tau))$$

- Variation of $|\Phi(\boldsymbol{\theta}(\tau))\rangle$ w.r.t. to parameters $\partial\theta$:

$$|\Phi(\boldsymbol{\theta}(\tau+\delta\tau))\rangle \approx |\Phi(\boldsymbol{\theta}(\tau))\rangle + \sum_j \frac{\partial\,|\Phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_j}\frac{\partial\theta_j}{\partial\tau}\delta\tau$$

- McLachlan's variational principle

$$\delta\left|\left(\frac{\partial}{\partial\tau}+(\hat{H}-E_\tau)\right)|\Phi(\boldsymbol{\theta}(\tau))\rangle\right| = 0$$

Evolution of parameters: $\dot{\boldsymbol{\theta}} = \mathbf{A}^{-1}\mathbf{C}$, $A_{ij} = \frac{\partial\,\langle\Phi|}{\partial\theta_i}\frac{\partial\,|\Phi\rangle}{\partial\theta_j}$ $C_i = -\frac{\partial\,\langle\Phi|}{\partial\theta_i}\hat{H}\,|\Phi\rangle$

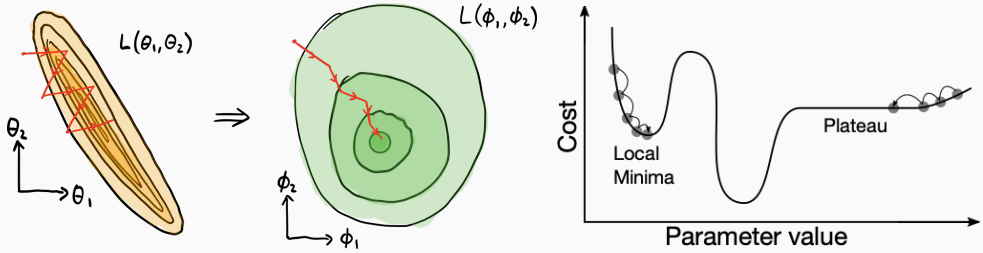
McLachlan, Molecular Physics, 8(1),39 (1964); McArdle *et al.*, npj Quantum, 5, 75 (2019); Yuan *et al.* Quantum 3, 191 (2019);

VarQITE – Connection to Quantum Natural Gradient

- Incorporate the geometry of the Ansatz, $|\Phi(\boldsymbol{\theta})\rangle = \hat{U}(\boldsymbol{\theta})\,|0\rangle$, with metric:

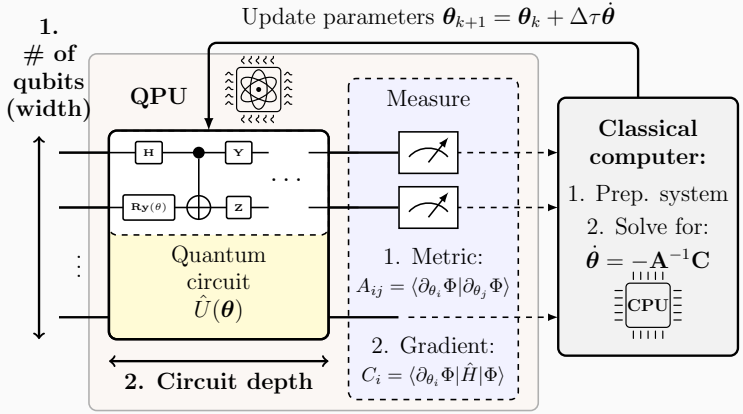
$$A_{ij} = \frac{\partial \langle \Phi|}{\partial \theta_i} \frac{\partial |\Phi\rangle}{\partial \theta_j}$$

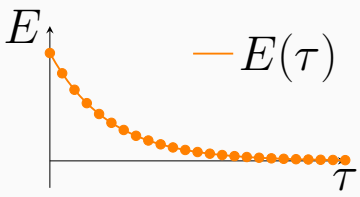
- *Natural gradient* update rule: $\dot{\boldsymbol{\theta}} = -\mathbf{A}^{-1}\boldsymbol{C}$



QITE Workflow

Can be performed in a NISQ-friendly hybrid approach





Sokolov, **WD**, Luo, Alavi, Tavernelli, Physical Review Research 5 (2), 023174 (2023)

VarQITE – Pros and Cons

Pros:

- No classical optimization
- Convergence (more) robust against noise
- Applicable to open/transport problems (**non Hermitian Hamiltonians**)

Cons:

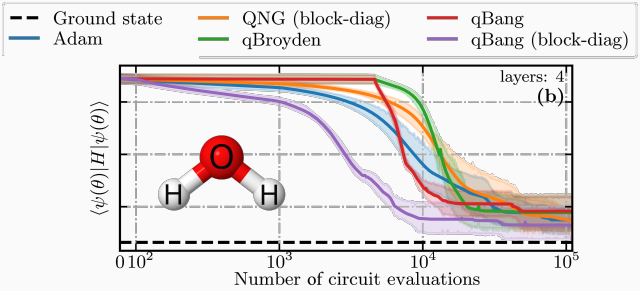
- 2^{nd} order method $\rightarrow$ metric $\mathbf{A}$ needs to be measured – n_θ^2 scaling with n_θ parameters
- $\mathbf{A}$ can be singular $\rightarrow$ inversion $\mathbf{A}^{-1}$ can problematic

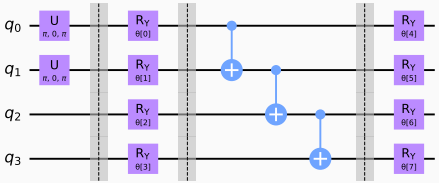
Improving classical optimization – qBANG – Quantum 8, 1313

Quasi-Newton approximation and an adaptive momentum (ADAM) approach to update the metric (with Sherman-Morrison formula direct update of $\mathbf{A}^{-1}$!)

$$\mathbf{A}_{k+1} \approx (1-\epsilon_k)\mathbf{A}_k + \epsilon_k \mathbf{C}_k \mathbf{C}_k^T$$

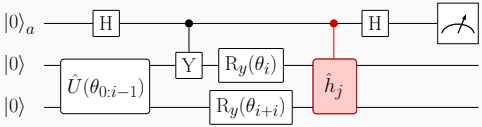
Immense reduction in circuit evaluations and improved convergence





VarQITE with non-Hermitian $\bar{H}$

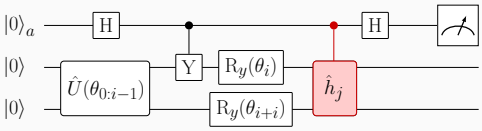
Hermitian gradient (i.e. lin. comb. of unitaries): $C_i = \frac{\partial \langle \Phi |}{\theta_i} \hat{H} |\Phi\rangle, \quad \hat{H} = \sum_j c_j \hat{h}_j$



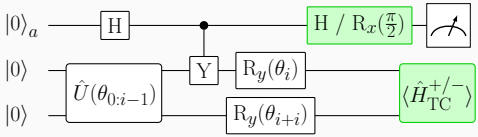
For TC: McArdle and Tew, arXiv:2006.11181

VarQITE with non-Hermitian $\bar{H}$

Hermitian gradient (i.e. lin. comb. of unitaries): $C_i = \dfrac{\partial \langle\Phi|}{\theta_i} \hat{H} |\Phi\rangle\,, \quad \hat{H} = \sum_j c_j \hat{h}_j$



For TC: McArdle and Tew, arXiv:2006.11181



Sokolov, **WD**, et al, PR Res. 5 (2), 023174

In the TC case: split non-Hermitian Hamiltonian in Hermitian and anti-Hermitian part:

$$\hat{H}^{+}_{TC} = \hat{H}_{TC} + \hat{H}^{\dagger}_{TC}, \qquad \hat{H}^{-}_{TC} = \hat{H}_{TC} - \hat{H}^{\dagger}_{TC}$$

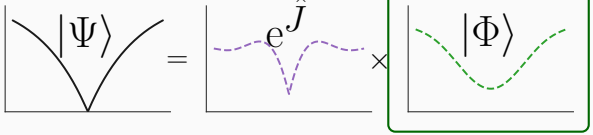
$$C_i = \frac{1}{2}\left(\langle\partial_{\theta_i}\Phi|\hat{H}_{TC}|\Phi\rangle + \langle\Phi|\hat{H}^{\dagger}_{TC}|\partial_{\theta_i}\Phi\rangle\right) = \frac{C^{+}_i + C^{-}_i}{4}$$

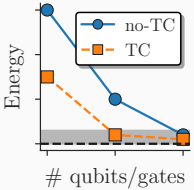
$$C^{+}_i = 2\langle\partial_{\theta_i}\Phi|\hat{H}^{+}_{TC}|\Phi\rangle, \qquad C^{-}_i = 2\langle\partial_{\theta_i}\Phi|\hat{H}^{-}_{TC}|\Phi\rangle$$

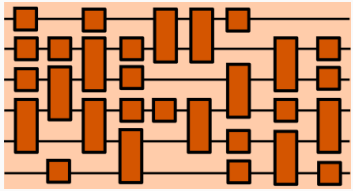
McArdle and Tew, arXiv:2006.11181 (2020); Sokolov, **WD**, Luo, Alavi, Tavernelli, Physical Review Research 5 (2), 023174 (2023) Schuld *et al.*, Phys. Rev. A 99, 032331 (2019)

Results: Hubbard model and ab initio quantum chemistry problems

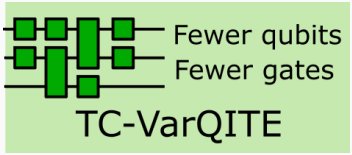
TC for QC – Motivation







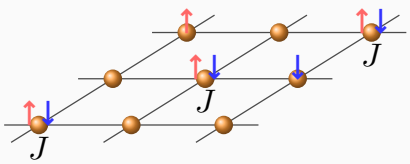
$$e^{-\hat{J}} \hat{H} e^{\hat{J}}$$

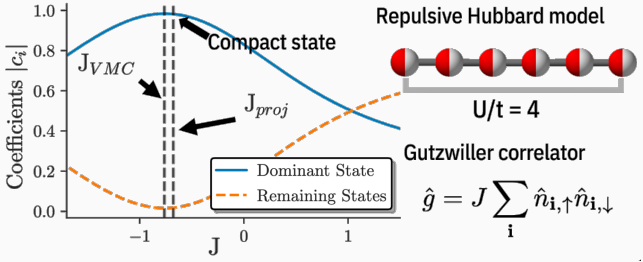


Reduce circuit depth with Transcorrelation

Suppress energetically unfavourable double occupancies via the *Gutzwiller* Ansatz:

$$\hat{g} = J \sum_i \hat{n}_{i\uparrow}\hat{n}_{i\downarrow} : \qquad \hat{H} \quad \rightarrow \quad \mathrm{e}^{-\hat{g}}\,\hat{H}\,\mathrm{e}^{\hat{g}}$$



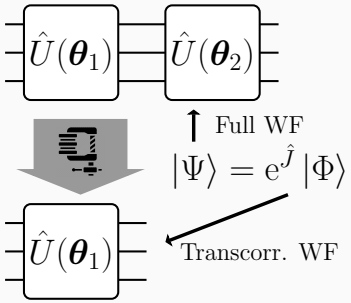


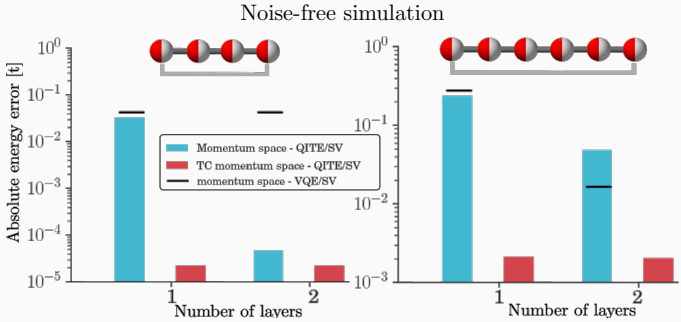
- **Increased compactness** of the right EV, due to downfolding of correlations into Hamiltonian
- Does the increased compactness/more single reference character have an impact on the necessary quantum Ansatz depth?

Gutzwiller, PRL 10, **159** (1963); Tsuneyuki, Prog. Theor. Phys. Supp., **176**, 134 (2008); Scuseria et al., PRB, **91**, 041114 (2015); **WD**, Luo, Alavi, PRB, **99**, 075119 (2019); Sokolov, **WD**, Luo, Alavi, Tavernelli, Physical Review Research 5 (2), 023174 (2023)

Results – Hubbard model – Phys. Rev. Research 5 (2), 023174

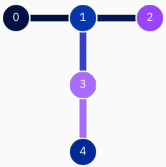
Transcorrelation ⇒ **shallower quantum circuits** for accurate results!

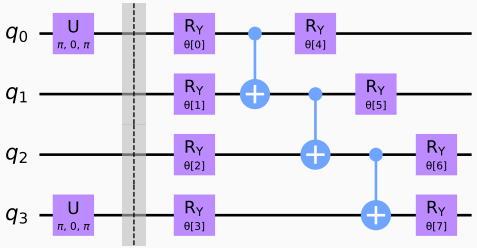




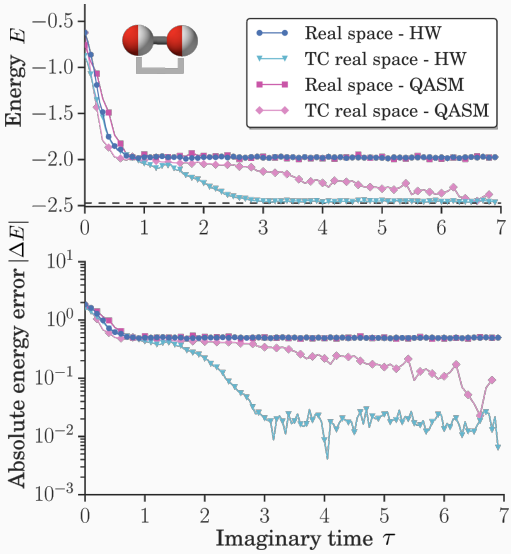
Noiseless statevector results, UCCSD Ansatz

Experimental results for the Hubbard model on ibmq_lima





Hardware-efficient RY Ansatz

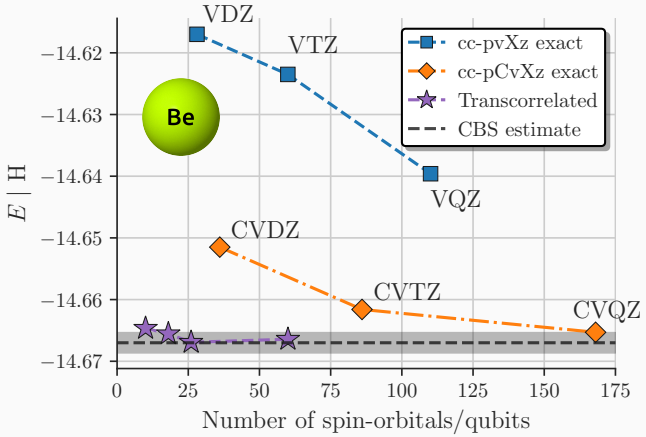


Sokolov, **WD**, Luo, Alavi, Tavernelli, Physical Review Research 5 (2), 023174 (2023)

Beryllium atom – JCTC 20 (10), 4146

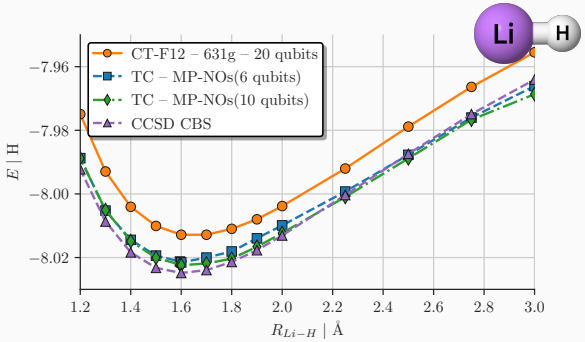
Beryllium atom – exact simulation of a quantum device (no noise)

Goal: complete basis set (CBS) limit → full description to compare with experiment

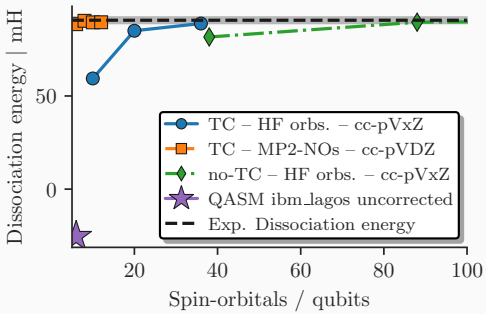


Lithium hydride – LiH – JCTC 20 (10), 4146

Lithium hydride – Potential energy surface and dissociation energy



Energy vs. bond distance – noiseless simulation



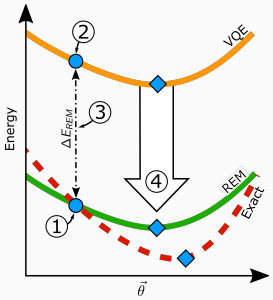
Noiseless and simulation of noisy quantum device (QASM) with hardware-efficient Ansatz

*CT-F12: Motta *et al.*, Phys. Chem. Chem. Phys. **22**, 24270, 2020; †Haeffler *et al.*, Phys. Rev. A, 1996, 53, 6, 4127 (1996); **WD**, Sokolov, Liao, Lopez Rios, Rahm, Alavi, Tavernelli, JCTC 20 (10), 4146

Reference-state Error Mitigation

Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry, *J. Chem. Theory Comput.*, **19**, 3, 783 (2023)

P. Lolur, M. Skogh, **W. Dobrautz**, C. Warren, J. Biznárová, A. Osman, G. Wendin, J. Bylander, M. Rahm

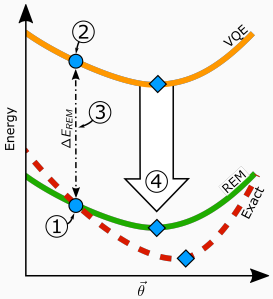


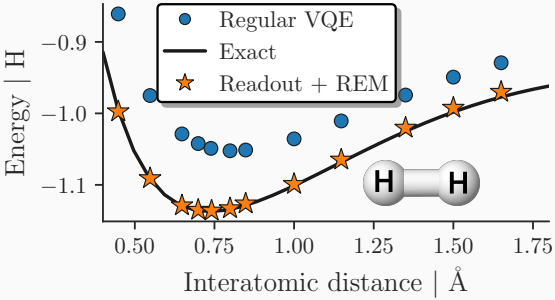
- Classically compute exact reference energy (i.e. Hartree-Fock), $E(\boldsymbol{\theta}_{\mathrm{ref}})$
- Measure reference energy on noisy device $\mathcal{E}(\boldsymbol{\theta}_{\mathrm{ref}})$, with reference parameters $\boldsymbol{\theta}_{\mathrm{ref}}$
- Calculate REM correction: $\Delta E_{\mathrm{REM}} = \mathcal{E}(\boldsymbol{\theta}_{\mathrm{ref}}) - E(\boldsymbol{\theta}_{\mathrm{ref}})$
- Correct final VQE energy with REM correction

Reference-state Error Mitigation

Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry, *J. Chem. Theory Comput.*, **19**, 3, 783 (2023)

P. Lolur, M. Skogh, **W. Dobrautz**, C. Warren, J. Biznárová, A. Osman, G. Wendin, J. Bylander, M. Rahm



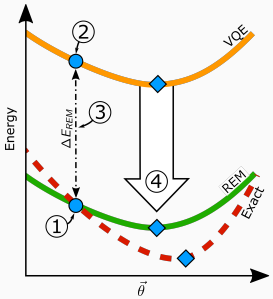


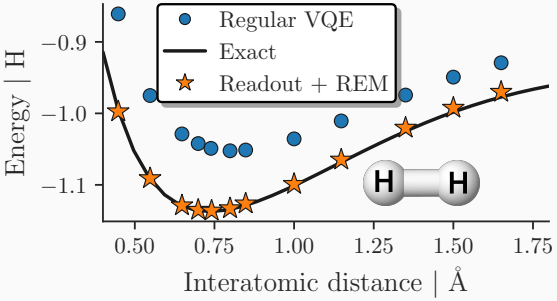
*Lolur, Skogh, **WD**, Warren, Biznárová, Osman, Tancredi, Wendin, Bylander, and Rahm, J. Chem. Theory Comput. 2023, **19**, 3, 783

Reference-state Error Mitigation

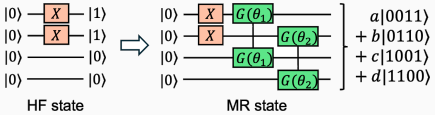
Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry, *J. Chem. Theory Comput.*, **19**, 3, 783 (2023)

P. Lolur, M. Skogh, **W. Dobrautz**, C. Warren, J. Biznárová, A. Osman, G. Wendin, J. Bylander, M. Rahm





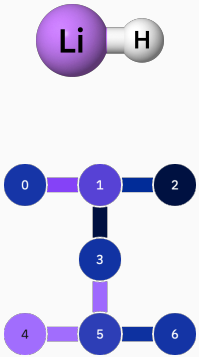
– **Multireference-state error mitigation for strong correlation**, H. Zou, E. Magnusson, H. Brunander, M. Rahm, **W. Dobrautz**, *to be submitted*

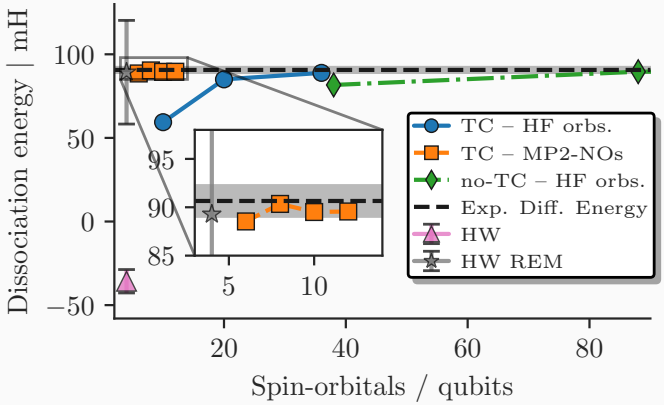


*Lolur, Skogh, **WD**, Warren, Biznárová, Osman, Tancredi, Wendin, Bylander, and Rahm, J. Chem. Theory Comput. 2023, **19**, 3, 783

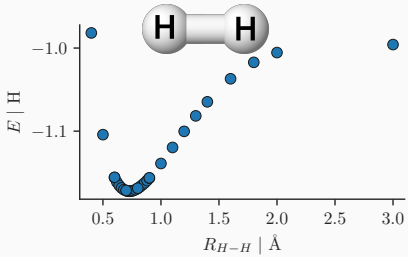
LiH – Experiment on IBM Quantum devices – JCTC 20 (10), 4146

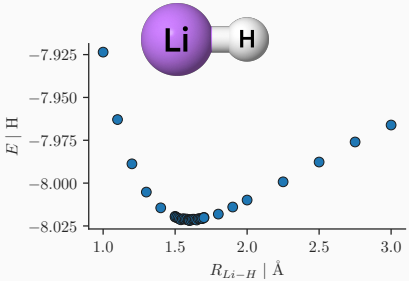
Hardware (HW) experiment: **lithium hydride** dissociation energy on `ibm_lagos`.
Hardware efficient RY Ansatz with linear entangling layer and parity encoding





Spectroscopic Constants

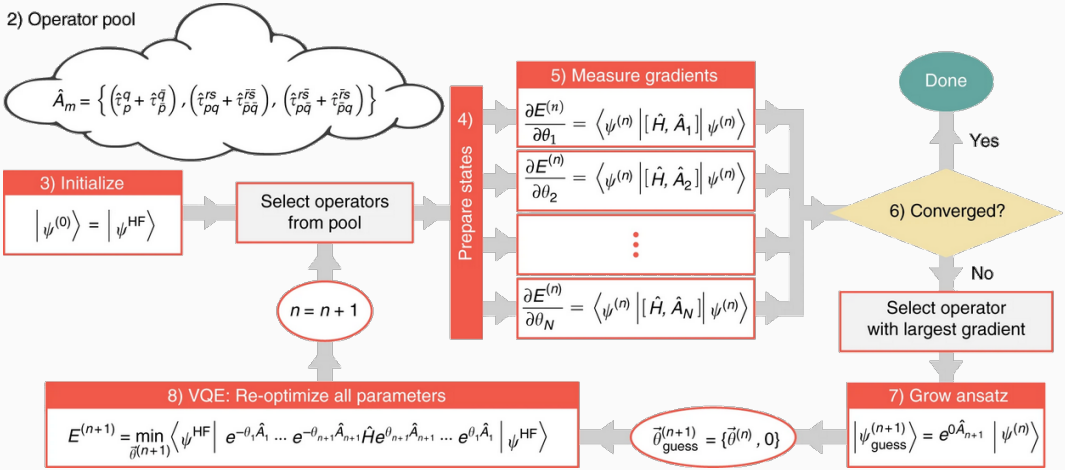




	H_2				LiH			
	qubits	R_e(Å)	D_0(eV)	ω_e(cm^{-1})	qubits	R_e(Å)	D_0(eV)	ω_e(cm^{-1})
no-TC	4	0.7330	3.67	4954	12	1.5422	2.66	1690
	8	0.7462	3.87	4297	22	1.6717	1.80	1283
	20	0.7609	4.19	4353	38	1.6154	2.17	1360
TC	**4**	**0.7428**	**4.69**	**4435**	**6**	**1.5998**	**2.47**	**1390**
Exp.		**0.7414**	**4.52**	**4401**		**1.5949**	**2.47**	**1406**

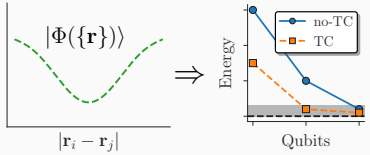
Adaptive Circuit Ansätze

Adaptive Circuit Ansätze

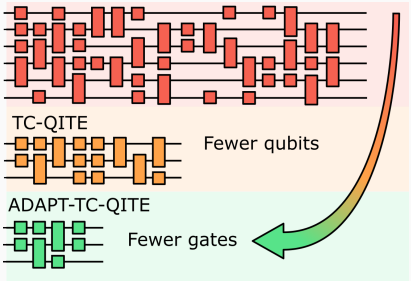


From: Grimsley et al. Nat Commun 10, 3007

Transcorrelation with Adaptive Circuit Ansätze



Smaller basis → fewer qubits



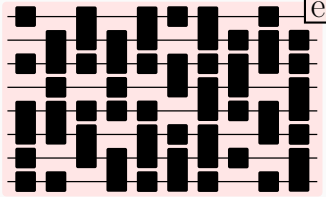
Full operator pool:

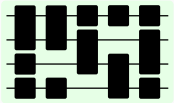
$\hat{U}_1, \hat{U}_2, \hat{U}_3, \hat{U}_4, \hat{U}_5, \hat{U}_6, \hat{U}_7, \hat{U}_8, \hat{U}_9, \hat{U}_{10}, \ldots$

ADAPT:

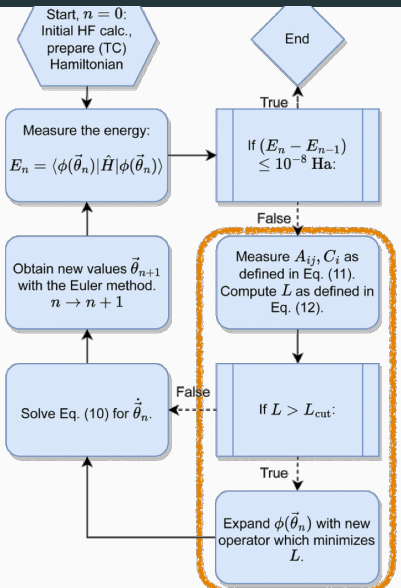
$\hat{U}'_1, \hat{U}'_2, \hat{U}'_3, \hat{U}'_4$

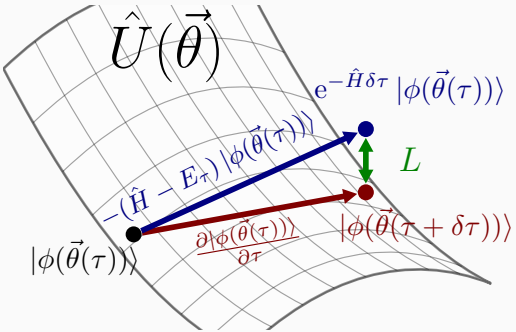
$e^{-\hat{J}} \hat{H} e^{\hat{J}}$





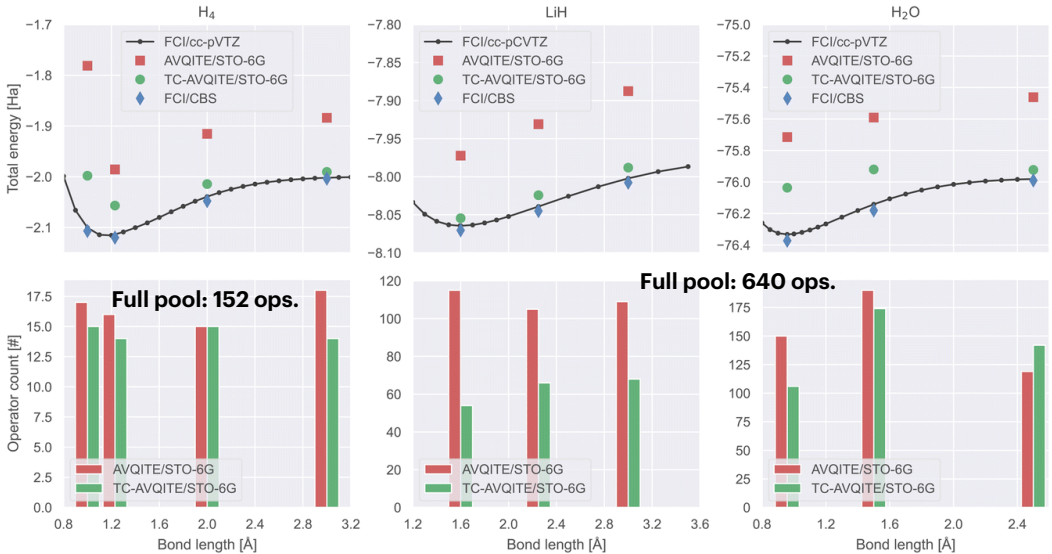
TC-Adaptive-VarQITE – Faraday Discuss., 254, 402





N. Gomes et al, Adv. Quantum Technol., 2021, 4, 2100114; Magnusson, Fitzpatrick, Knecht, Rahm and **WD**, Faraday Discuss., 254, 402 (2024)

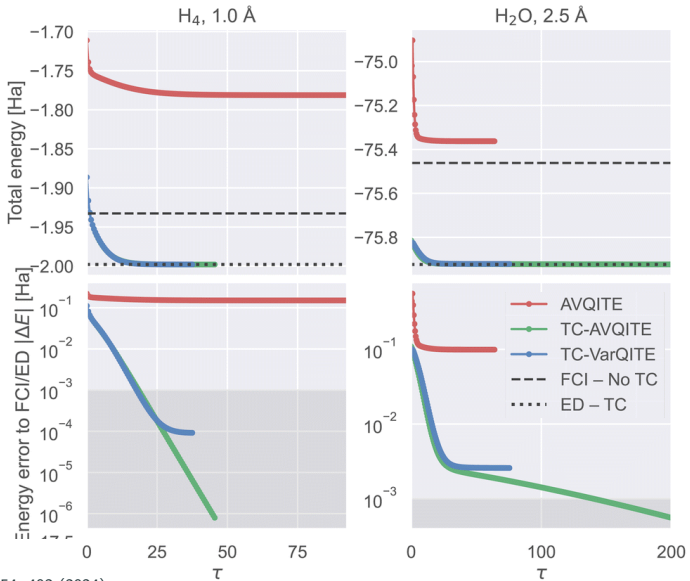
TC-AVQITE – Results – Qubit-ADAPT UCCSD Pool



Magnusson, Fitzpatrick, Knecht, Rahm and **WD**, Faraday Discuss., 254, 402 (2024)

TC-AVQITE – Improved Convergence

- More compact right eigenvector → shallower circuits
- Better convergence of TC calculations compared to non-TC results
- Additionally improved convergence compared to "full" UCCSD Ansätze.

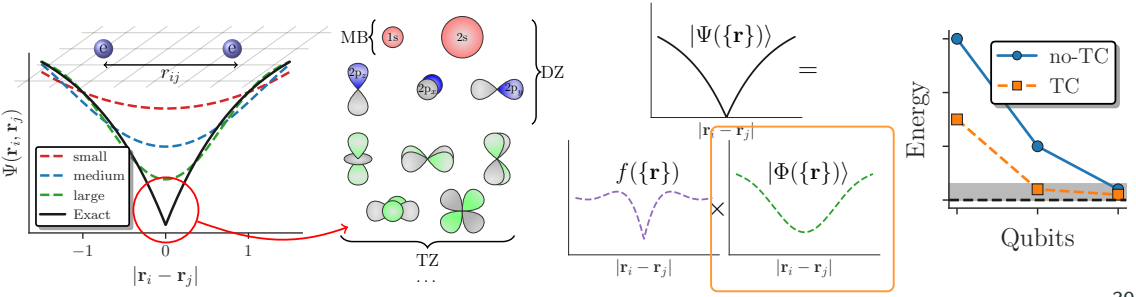


Magnusson, Fitzpatrick, Knecht, Rahm and **WD**, Faraday Discuss., 254, 402 (2024)

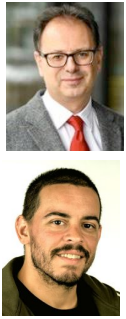
Conclusions

Conclusion – Transcorrelated Approach on Quantum Hardware

- The **TC method** partially transfers electronic correlations from the wavefunction into the Hamiltonian, **capturing the cusp condition**.
- **Reduce qubit requirements and circuit depth**, due to accurate results with a small basis sets and more compact (right) eigenvectors
- With efficient error mitigation techniques → **extends applicability of current and near-term quantum devices** to more relevant quantum chemistry problems.



Acknowledgments



Funding

SPONSORED BY THE
Federal Ministry of Education and Research

Group

PhDs

PostDocs

I AM HIRING!

Physics, Chemistry, Computer Science,
Quantum Technology, AI / Machine Learning, ...

Funding and Acknowledgements

Bundesministerium für Bildung und Forschung

SACHSEN

DRESDEN concept

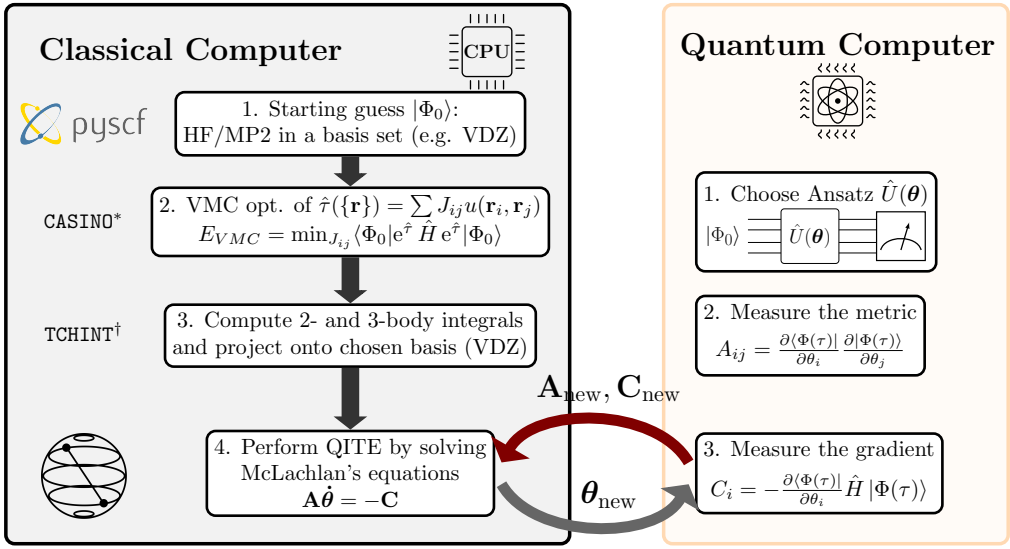
ScaDS.AI DRESDEN LEIPZIG

HELMHOLTZ HZDR

CASUS

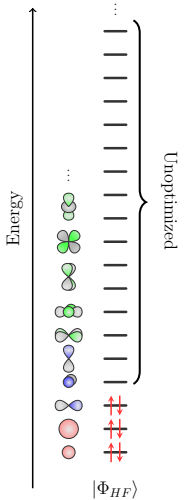
Thank you for your attention!

Workflow



*https://vallico.net/casinoqmc/ †https://gitlab.com/kguther/tchint ‡https://github.com/ghb24/NECI_STABLE

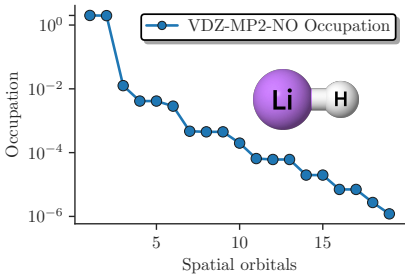
(Virtual) orbital optimization

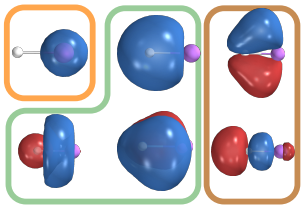


- "Standard basis sets" not optimized for the TC method

→ include effect of virtuals through orbital optimization / downfolding

→ e.g. natural orbitals (NO) from a "cheap" perturbation theory (MP2) calculation





FNO: Sosa *et al.*, Chem. Phys. Lett. **159** (2-3) 148 (1989); Taube and Bartlett Collect. Czech. Chem. Commun. **70**, 837 (2005); **NISQ:** Gonthier *et al.* arXiv:2012.04001; (2020); Verma *et al.* JCP **155**, 034110 (2021)

VarQITE

Three ingredients:

- Evolution of $|\Phi(\boldsymbol{\theta}(\tau))\rangle$ according to imaginary time Schrödinger equation, $|\Psi(\tau)\rangle = \mathrm{e}^{-\tau(\hat{H}-S_\tau)}\,|\Psi(0)\rangle$ for small $\delta\tau$: $\mathrm{e}^{-\delta\tau(\hat{H}-S_\tau)} \approx (1-\delta\tau(\hat{H}-S_\tau))$

$$|\Phi(\boldsymbol{\theta}(\tau+\delta\tau))\rangle \approx \Big[1-\delta\tau(\hat{H}-S_\tau)\Big]\,|\Phi(\boldsymbol{\theta}(\tau))\rangle \tag{1}$$

- Variation of $|\Phi(\boldsymbol{\theta}(\tau))\rangle$ w.r.t. to parameters $\partial\theta$:

$$|\Phi(\boldsymbol{\theta}(\tau+\delta\tau))\rangle \approx |\Phi(\boldsymbol{\theta}(\tau))\rangle + \sum_j \frac{\partial\,|\Phi(\boldsymbol{\theta}(\tau))\rangle}{\partial\theta_j}\frac{\partial\theta_j}{\partial\tau}\delta\tau \tag{2}$$

Equate r.h.s. of Eqs. (1) and (2) $\Rightarrow$

McLachlan's variational principle

$$\implies \sum_j \frac{\partial \left|\Phi(\boldsymbol{\theta}(\tau))\right\rangle}{\partial \theta_j} \dot{\theta}_j \approx -(\hat{H} - S_\tau) \left|\Phi(\boldsymbol{\theta}(\tau))\right\rangle, \quad \text{with} \quad \dot{\theta}_j = \frac{\partial \theta_j}{\partial \tau} \tag{3}$$

3. McLachlan's variational principle to minimize the distance between l.h.s and r.h.s. of (3)

$$\delta \left|\left| \left(\frac{\partial}{\partial \tau} + \hat{H} - S_\tau \right) \left|\Phi(\boldsymbol{\theta}(\tau))\right\rangle \right|\right| = 0, \quad \text{with} \quad \left|\left| \left|\Phi\right\rangle \right|\right| = \sqrt{\langle \Phi \,|\, \Phi \rangle}.$$

After some calculations we find a formula to update the parameters $\boldsymbol{\theta}$ to emulate imaginary time evolution on quantum computers

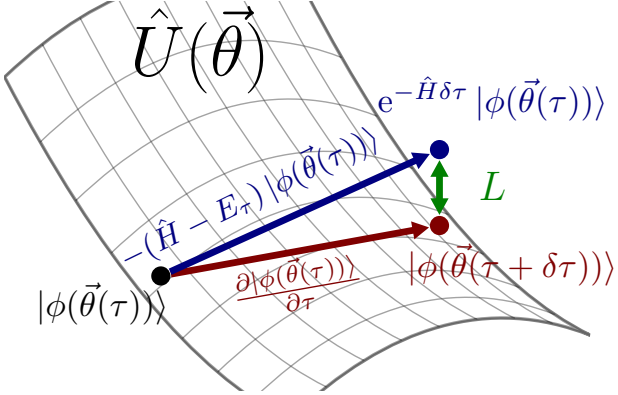
McLachlan, Molecular Physics, 8(1),39 (1964)

VarQITE

Simulate the imaginary-time (Wick-rotated) Schrödinger equation

$$\frac{\partial |\Psi(\tau)\rangle}{\partial \tau} = -(\hat{H} - E_\tau)|\Psi(\tau)\rangle \quad \rightarrow |\Psi(\tau)\rangle = \mathrm{e}^{-\delta\tau(\hat{H}-E_\tau)}|\Psi(0)\rangle$$

by mapping on time evolution of parameters $\boldsymbol{\theta}(\tau)$



- Ansatz $|\Phi(\boldsymbol{\theta}(\tau))\rangle = \hat{U}(\boldsymbol{\theta}(\tau))|0\rangle$
- McLachlan's variational principle

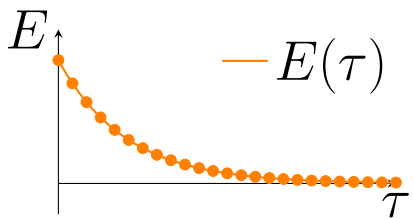
$$\delta \left| \left(\frac{\partial}{\partial \tau} + (\hat{H} - E_\tau) \right) |\Phi(\boldsymbol{\theta}(\tau))\rangle \right| = 0$$

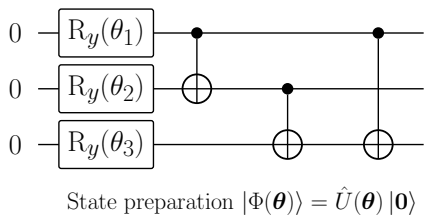
- Evolution of parameters $\mathbf{A} \cdot \dot{\boldsymbol{\theta}} = C$

$$A_{ij} = \frac{\partial \langle\Phi|}{\partial \theta_i} \frac{\partial |\Phi\rangle}{\partial \theta_j} \qquad C_i = -\frac{\partial \langle\Phi|}{\partial \theta_i} \hat{H} |\Phi\rangle$$

Quantum Imaginary Time Evolution

Imaginary-time evolution (ITE) applicable to efficiently solve for groundstate energy on quantum hardware





Quantum ITE: Map imaginary-time evolution of $|\Phi(\tau)\rangle$ to change of gate parameters $\boldsymbol{\theta}(\tau)$

$$\dot{\boldsymbol{\theta}} = \mathbf{A}^{-1}\boldsymbol{C}, \qquad A_{ij} = \frac{\partial \langle\Phi|}{\partial\theta_i}\frac{\partial |\Phi\rangle}{\partial\theta_j} \qquad C_i = -\frac{\partial \langle\Phi|}{\partial\theta_i}\hat{H}\,|\Phi\rangle$$

with the metric (quantum Fisher information), $\mathbf{A}$, and the gradient of the cost function, $\mathbf{C}$.

primo.ai; pennylane.ai; McArdle *et al.*, npj Quantum, 5, 75 (2019); Yuan *et al.* Quantum 3, 191 (2019);

Connection to natural gradient

Equation for change in parameters $\boldsymbol{\theta}$ due to McLachlan's variational principle to enable Ansatz-based quantum imaginary time evolution:

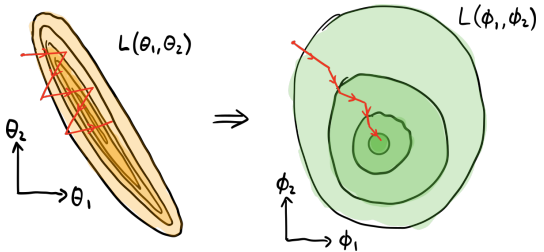
$$\sum_j A_{ij}\dot{\theta}_j = C_i, \quad \implies \dot{\boldsymbol{\theta}} = \mathbf{A}^{-1}\boldsymbol{C}$$

with the metric:

$$A_{ij} = \frac{\partial \langle\Phi|}{\partial\theta_i}\frac{\partial |\Phi\rangle}{\partial\theta_j}$$

and energy gradient:

$$C_i = -\frac{\partial \langle\Phi|}{\partial\theta_i}\hat{H}\,|\Phi\rangle$$

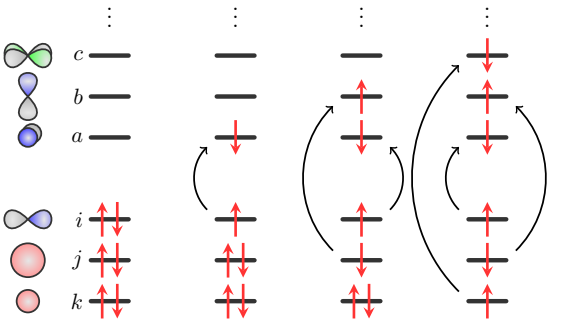


* Motta *et al.*, Nature Physics, 16, 205 (2020); pennylane.ai

Exponential scaling of Full Configuration Interaction

FCI $\Rightarrow$ exact solution in a given basis: linear combination of determinants

$$|\Psi\rangle = |\Phi_{HF}\rangle + \sum_i c_i |\Phi_i\rangle$$



All possible excitations from HF determinant

Number of possible states for given number of electrons, N, and orbitals, n, $\sim \binom{N}{n}$

Mol.	#orbitals	#electrons	#states
H_2	2	2	4
LiH	4	4	36
Be_2	8	8	4900
H_2O	12	12	$\sim 8 \cdot 10^5$
C_2H_4	16	16	$\sim 16 \cdot 10^6$
F_2	18	18	$\sim 2 \cdot 10^9$

Variational Quantum Monte Carlo to optimize Jastrow factors

Minimize variational energy, by optimizing trial wavefunction parameters J_{ij}:

$$E_{VMC} = \min_{\hat{J}(J_{ij})} \frac{\langle \Phi_0 | \mathrm{e}^{\hat{J}} \hat{H} \, \mathrm{e}^{\hat{J}} | \Phi_0 \rangle}{\langle \Phi_0 | \mathrm{e}^{2\hat{J}} | \Phi_0 \rangle}, \quad |\Phi_T\rangle = \mathrm{e}^{\hat{J}} |\Phi_0\rangle$$

- The choice of trial wavefunction is critical in VMC calculations $\rightarrow$ accuracy limited by $|\Phi_T\rangle = \mathrm{e}^{\hat{J}} |\Phi_0\rangle$!
- Hartree-Fock state usually first starting point for $|\Phi_0\rangle$, but more elaborate/accurate states possible...
- Polynomial scaling $\sim N^3$
- Such a VMC calculations to optimize J_{ij} with a HF state $|\Phi_0\rangle = |\Phi_{HF}\rangle$ our **starting point** for the **transcorrelated method**

Jastrow s.t. Hamiltonian in 2nd quantised form

$$\bar{H} = \sum_{pq,\sigma} h_q^p a_{p,\sigma}^\dagger a_{q,\sigma} + \frac{1}{2} \sum_{pqrs} (V_{rs}^{pq} - K_{rs}^{pq}) \sum_{\sigma,\tau} a_{p,\sigma}^\dagger a_{q,\tau}^\dagger a_{s,\tau} a_{r,\sigma}$$
$$- \frac{1}{6} \sum_{pqrstu} L_{stu}^{pqr} \sum_{\sigma\tau\lambda} a_{p,\sigma}^\dagger a_{q,\tau}^\dagger a_{r,\lambda}^\dagger a_{u,\lambda} a_{t,\tau} a_{s,\sigma}$$

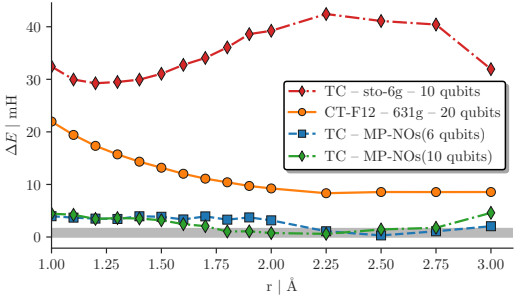
with

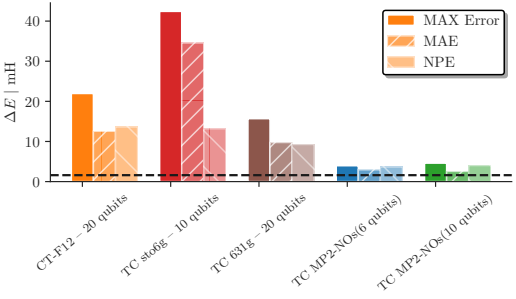
$$K_{rs}^{pq} = \langle \phi_p \phi_q | \hat{K} | \phi_r \phi_s \rangle$$
$$L_{stu}^{pqr} = \langle \phi_p \phi_q \phi_r | \hat{L} | \phi_s \phi_t \phi_u \rangle \quad \text{(48-fold symmetry in } L \text{ for real orbitals)}$$

Both integrals K and L are computed numerically using standard DFT grids over gaussian orbitals. The main problem is the storage of L. Current limit $\approx$ 80 orbitals

Cohen, Luo, Guther, **WD**, Tew, Alavi, JCP **151** (6), 061101 (2019);

LiH


40
30
20
10
0
ΔE | mH
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
r | Å
TC – sto-6g – 10 qubits
CT-F12 – 631g – 20 qubits
TC – MP-NOs(6 qubits)
TC – MP-NOs(10 qubits)

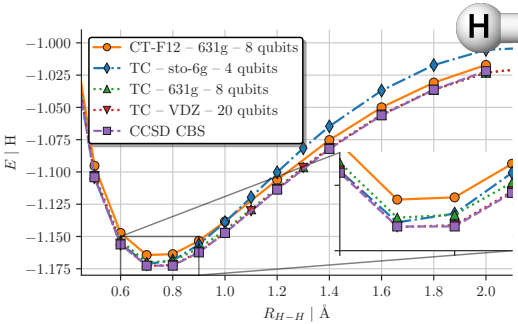

40
30
20
10
0
ΔE | mH
MAX Error
MAE
NPE
CT-F12 – 20 qubits
TC sto6g – 10 qubits
TC 631g – 20 qubits
TC MP2-NOs(6 qubits)
TC MP2-NOs(10 qubits)

Hydrogen molecule

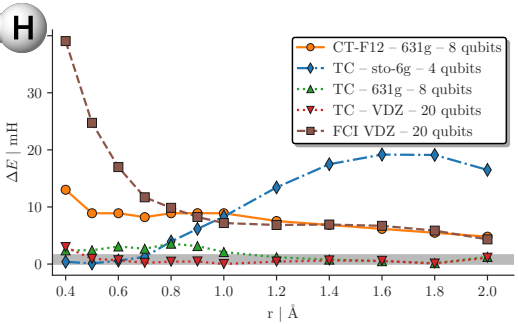
Favorite quantum chemistry test case: Hydrogen molecule – H_2

CT-F12 approximated explicitly correlated method, by Motta *et al.*[*]

Exact statevector simulation – UCCSD Ansatz



Energy vs. bond distance



Error wrt. CBS result vs. bond distance

[*] CT-F12: Motta *et al.*, Phys. Chem. Chem. Phys. **22**, 24270, 2020

Imaginary Time Evolution (ITE) – Normalization

$E_0 - S_\tau = 0$ would require knowledge of ground state. Alternative, and also to ensure proper normalization:

For small time-steps $\Delta\tau$ approximate exponential by first-order Taylor approximation* and obtain iterative solution:

$$\mathrm{e}^{-\Delta\tau(\hat{H}-S_\tau)} \approx 1 - \Delta\tau(\hat{H} - S_\tau) + \mathcal{O}(\Delta\tau^2) \quad \rightarrow \quad |\Psi(\tau+\Delta\tau)\rangle = \Big[1 - \Delta\tau(\hat{H} - S_\tau)\Big]\,|\Psi(\tau)\rangle \tag{4}$$

Assuming $\langle\Psi(\tau)\,|\,\Psi(\tau)\rangle = 1$:

$$\begin{aligned}\langle\Psi(\tau+\Delta\tau)\,|\,\Psi(\tau+\Delta\tau)\rangle &= \langle\Psi(\tau)|\Big[1 - \Delta\tau(\hat{H} - S_\tau)\Big]^2|\Psi(\tau)\rangle \stackrel{!}{=} 1 \\ &= \underbrace{\langle\Psi(\tau)\,|\,\Psi(\tau)\rangle}_{=1} - 2\Delta\tau \underbrace{\Big(\langle\Psi(\tau)|\hat{H}|\Psi(\tau)\rangle - S_\tau\Big)}_{\stackrel{!}{=}0} + \mathcal{O}(\Delta\tau^2)\end{aligned}$$

$\Longrightarrow S_\tau = \langle\Psi(\tau)|\hat{H}|\Psi(\tau)\rangle$ ensures normalization of $|\Psi(\tau)\rangle$ and allows ITE to converge to the groundstate

* for $\Delta\tau < 1/E_W$, with $E_W = E_{max} - E_0$ being the many-body spectral width, Trivedi and Ceperley, Phys. Rev. B, 41, 4552 (1990)

$$\langle\Phi|\Big(\partial/\partial\tau+\hat{H}-S_\tau\Big)^\dagger\Big(\partial/\partial\tau+\hat{H}-S_\tau\Big)|\Phi\rangle=\sum_{ij}\frac{\partial\langle\Phi|}{\partial\theta_i}\frac{\partial|\Phi\rangle}{\partial\theta_j}\dot{\theta}_i\dot{\theta}_j+\sum_i\frac{\partial\langle\Phi|}{\partial\theta_i}(\hat{H}-S_\tau)|\Phi\rangle\,\dot{\theta}_i$$

$$+\sum_i\langle\Phi|(\hat{H}-S_\tau)\frac{\partial|\Phi\rangle}{\partial\theta_i}\dot{\theta}_i+\langle\Phi|(\hat{H}-S_\tau)^2|\Phi\rangle \tag{5}$$

Variations in $\dot{\theta}_i \Longrightarrow$ and focusing on one term in $\dot{\theta}_i$:

$$\frac{\partial||(\partial/\partial\tau+\hat{H}-S_\tau)|\Phi\rangle||}{\partial\dot{\theta}_i}=\sum_j\left(\frac{\partial\langle\Phi|}{\partial\theta_i}\frac{\partial|\Phi\rangle}{\partial\theta_j}+\frac{\partial\langle\Phi|}{\partial\theta_j}\frac{\partial|\Phi\rangle}{\partial\theta_i}\right)\dot{\theta}_j$$

$$+\frac{\partial\langle\Phi|}{\partial\theta_i}\hat{H}|\Phi\rangle+\langle\Phi|\hat{H}\frac{\partial|\Phi\rangle}{\partial\theta_i}-S_\tau\left(\frac{\partial\langle\Phi|}{\partial\theta_i}|\Phi\rangle+\langle\Phi|\frac{\partial|\Phi\rangle}{\partial\theta_i}\right)$$

with

$$\langle\Phi\,|\,\Phi\rangle=1,\quad\rightarrow\quad\frac{\partial\langle\Phi\,|\,\Phi\rangle}{\partial\theta_i}=\frac{\partial\langle\Phi|}{\partial\theta_i}|\Phi\rangle+\langle\Phi|\frac{\partial|\Phi\rangle}{\partial\theta_i}=0$$

$$\Longrightarrow\frac{\partial||(\partial/\partial\tau+\hat{H}-S_\tau)|\Phi\rangle||}{\partial\dot{\theta}_i}=\sum_j A_{ij}\dot{\theta}_j-C_i\stackrel{!}{=}0$$

Monotonic energy convergence

$$\frac{dE(\tau)}{d\tau} = \langle\Phi(\tau)|\,\hat{H}\,\frac{d\,|\Phi(\tau)\rangle}{d\tau} = \sum_i \langle\Phi(\tau)|\,\hat{H}\,\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_i}\dot{\theta}_i \tag{6}$$

$$= -\sum_i C_i\dot{\theta}_i = -\sum_i C_i A^{-1}_{ij} C_j \leq 0,$$

if $\mathbf{A}^{-1}$ is positive. For arbitrary $\mathbf{x} = (x_1, x_2, \ldots, x_N)^T$:

$$\mathbf{x}^\dagger \cdot \mathbf{A} \cdot \mathbf{x} = \sum_{ij} x_i^* A_{ij} x_j = \sum_{ij} x_i^* \left(\frac{\partial\,\langle\Phi(\tau)|}{\partial\theta_i}\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_j} \right) x_j$$

with a general $|\zeta\rangle = \sum_i x_i \frac{\partial|\Phi(\tau)\rangle}{\partial\theta_i}$:

$$\sum_{ij} x_i^* \left(\frac{\partial\,\langle\Phi(\tau)|}{\partial\theta_i}\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_j} \right) x_j = \langle\zeta|\zeta\rangle \geq 0.$$

And we only consider non-zero eigenvalues in case $\mathbf{A}^{-1}$ is singular

Linear combination of unitaries

Assuming: each unitary gate depends only on one parameter θ_i and each U_i is a rotation or controlled rotation gate.

$$\frac{\partial U_i(\theta_i)}{\partial \theta_i} = \sum_k f_{k,i} U_i(\theta_i) u_{k,i}, \quad \rightarrow \quad \frac{\partial \left|\Phi(\tau)\right\rangle}{\partial \theta_i} = \sum_k f_{k,i} V'_{k,i} \left|\mathbf{0}\right\rangle \tag{7}$$

with a unitary operator $u_{k,i}$, scalar parameter $f_{k,i}$ and:

$$V'_{k,i} = U_n(\theta_n) \dots U_i(\theta_i) u_{k,i} \dots U_1(\theta_1)$$

Evaluation $\mathbf{A}$ and C with Quantum Circuits

How do we measure the metric, $\mathbf{A}$, and the gradient, $\mathbf{C}$ on quantum hardware.

$$A_{ij} = \frac{\partial \langle \Phi |}{\partial \theta_i} \frac{\partial | \Phi \rangle}{\partial \theta_j}, \qquad C_i = -\frac{\partial \langle \Phi |}{\partial \theta_i} \hat{H} | \Phi \rangle$$

$|\Phi(\boldsymbol{\theta}(\tau))\rangle$ encoded by unitary gates acting on initial state: Ansatz $\hat{U}(\boldsymbol{\theta}(\tau))$

$$|\Phi(\boldsymbol{\theta}(\tau))\rangle = \hat{U}_n(\theta_n(\tau)) \cdots \hat{U}_i(\theta_i(\tau)) \cdots \hat{U}_1(\theta_1(\tau)) |\mathbf{0}\rangle = \hat{U}(\boldsymbol{\theta}(\tau)) |\mathbf{0}\rangle$$

- Numerical differentiation/approximation: $\frac{\partial \hat{U}_i(\theta_i)}{\partial \theta_i} \approx \frac{\hat{U}_i(\theta_i + \Delta\theta_i) - \hat{U}_i(\theta_i)}{\Delta \theta_i}$
- Parameter-shift rule* (for single qubit gates): $R_z(\theta_i) = \mathrm{e}^{-i\theta_i \sigma_z} \rightarrow \frac{\partial U_i(\theta_i)}{\partial \theta_i} = -i\sigma_z R_Z(\theta_i)$
- Linear combination of unitaries* (for general gates), see (7)ff for details

*Schuld *et al.*, Phys. Rev. A 99, 032331 (2019); Romero *et al.*, Quantum Science and Technology, 4, 1 (2019); Li and Benjamin, Phys. Rev. X 7, 021050 (2017);

Derivative Example

- $U_i(\theta_i)$ is a single qubit rotation: $R_Z(\theta_i) = \mathrm{e}^{-i\theta_i\sigma_z}$:

$$\frac{\partial U_i(\theta_i)}{\partial \theta_i} = -\frac{i}{2}\sigma_z R_Z(\theta_i)$$

  $\rightarrow$ add an extra σ_z gate with factor $-i/2$

- $U_i(\theta_i)$ a controlled rotation: $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes R_Z(\theta_i)$:

$$\frac{\partial U_i(\theta_i)}{\partial \theta_i} = |1\rangle\langle 1| \otimes \partial R_Z(\theta_i)/\partial\theta_i = -\frac{i}{2}|1\rangle\langle 1| \otimes \sigma_z R_Z(\theta_i)$$

  $\rightarrow$ realized with

$$u_{1,i} = I \otimes \sigma_z, f_{1,i} = -i/4,$$
$$u_{2,i} = \sigma_z \otimes \sigma_z, f_{2,i} = i/4$$

  in $\sum_k f_{k,i} U_i(\theta_i) u_{k,i}$ from previous slide

Evaluation $\mathbf{A}$ and $\mathbf{C}$ with Quantum Circuits – cont.

$$A_{ij} = \sum_{k,l} f^*_{k,i} f_{l,j} \langle \mathbf{0} | V^{\prime\dagger}_{k,i} V^{\prime}_{l,j} | \mathbf{0} \rangle, \qquad C_i = \sum_{k,l} f^*_{k,l} \lambda_l \langle \mathbf{0} | V^{\prime\dagger} h_l \hat{V} | \mathbf{0} \rangle \tag{8}$$

with $\hat{H} = \sum_l \lambda_l h_l$. Both $\mathbf{A}$ and $\boldsymbol{C}$ are of the form $a \cdot \mathrm{e}^{i\phi} \langle \mathbf{0} | \hat{U} | \mathbf{0} \rangle$ and can be evaluated on a quantum circuit.