Reducing necessary quantum hardware resources with explicitly correlated methods

Werner Dobrautz
Chalmers University of Technology

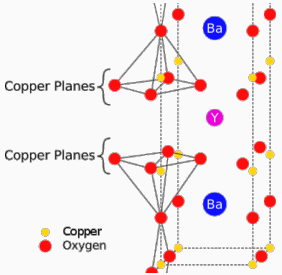
Algorithmiq, April 22, 2023

Outline

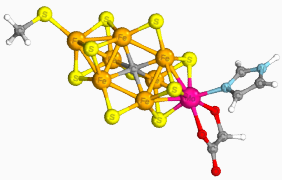
- Electronic Structure Theory – Quantum Chemistry
- Correlated Ansatz to reduce the computational footprint on quantum hardware
- Intermezzo: Quantum Imaginary Time Evolution – QITE
- Applications: Circuit depth – Hubbard model
- Applications: Circuit width – *Ab initio* problems
- Conclusions and Outlook

Electronic Structure Theory – Quantum Chemistry

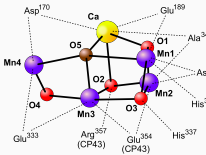
Ab Initio Quantum Chemistry – Electronic Structure Theory



YBCO: Unconventional high-T_c superconductivity



FeMoCo: primary cofactor of nitrogenase → nitrogen fixation



Manganese-Cadmium-Oxygen Clusters: Oxygen evolving clusters in photosystem II

Surprisingly small systems responsible for interesting physical/chemical properties!

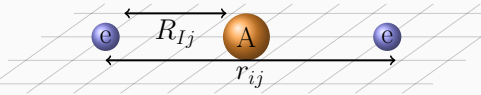
Strongly correlated ⇒ challenging systems for computational approaches!

Accurate theoretical understanding at nano-scale for bottom-up materials design!

Ab Initio Quantum Chemistry – Electronic Structure Theory

All necessary information of a quantum system contained in electronic **molecular Hamiltonian** (Born-Oppenheimer approx., atomic units and first quantization)

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



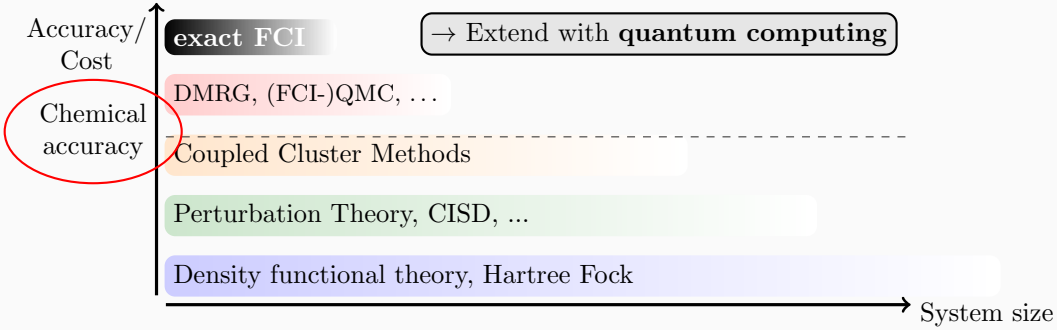
Electronic properties: Ground- and low-lying excited state properties, energy differences, polarization, response functions, ...

Task: Solve the Schrödinger equation derived from first principles

$$\hat{H}\,|\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n)\rangle = E\,|\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n)\rangle$$

Target: High / chemical accuracy to ensure predictability, interpretability and comparison with experimental results.

Hierarchy of methods

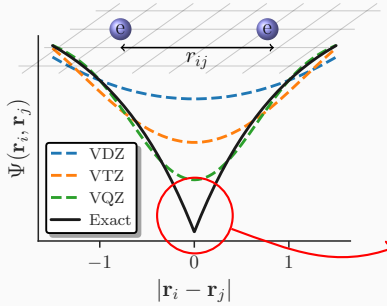


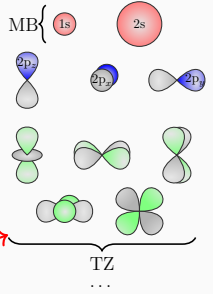
Highly accurate methods only applicable to **very small** system sizes. Current quantum computing calculations/experiments use small/**minimal basis sets** far from experimental results, due to **limited number of qubits**

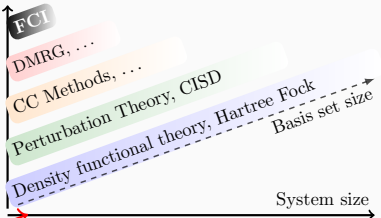
Problems for accurate description: Cusp condition

Cusp condition: Singularity of Coulomb potential, $\frac{1}{r_{ij}}$, for $r_{ij} = 0 \rightarrow$ sharp cusp of exact wavefunction $\Psi(\{\mathbf{r}\})$ at electron coalescence ($r_{ij} = 0$)

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





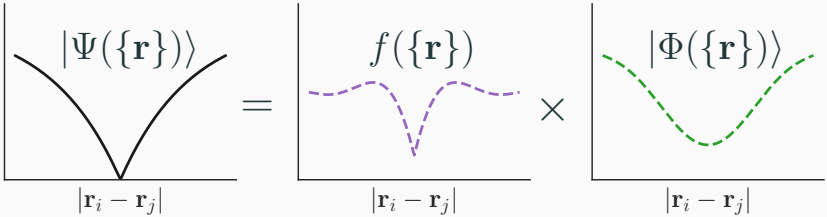


Kato, Communications on Pure and Applied Mathematics **10** (2), 151 (1957)

Correlated Ansatz to reduce the computational footprint on quantum hardware

Cusp Condition – Explicitly Correlated Ansatz

Short range behavior → **dynamic correlation**. Important for accurate results → Necessitates **large basis set expansion**.

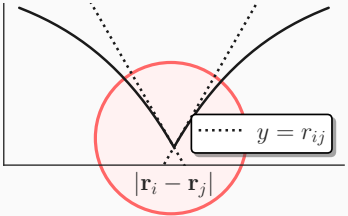


Describe the cusp exactly and capture part of correlation with a **correlated wavefunction Ansatz**

$$|\Psi(\{\mathbf{r}\})\rangle = f(\{\mathbf{r}\})\,|\Phi(\{\mathbf{r}\})\rangle$$

Hylleras, Z. Phys. 54, 347 (1929)

Explicitly Correlated methods



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

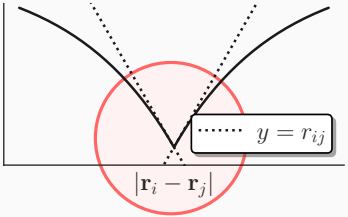
R12 methods*: $|\Psi\rangle = r_{ij}\,|\Phi\rangle$

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

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

* Kutzelnigg, Theoretica chimica acta 68, 445 (1985); † Ten-no, J. Chem. Phys. 121, 117 (2004); ‡ Jastrow, Phys. Rev. **98**, 1479 (1955);

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

$\exp(-x) \approx 1 - x + \mathcal{O}(x^2), \qquad \tilde{r}_{ij} = \dfrac{r_{ij}}{1+r_{ij}}, \quad \lim_{r_{ij}\to 0} \tilde{r}_{ij} \to 0, \ \lim_{r_{ij}\to\infty} \tilde{r}_{ij} \to 1$

* Kutzelnigg, Theoretica chimica acta 68, 445 (1985); † Ten-no, J. Chem. Phys. 121, 117 (2004); ‡ Jastrow, Phys. Rev. **98**, 1479 (1955);

Variational Quantum Monte Carlo to optimize Jastrow factors

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

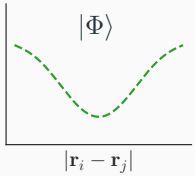
$$E_{VMC} = \min_{\hat{J}(J_{ij})} \frac{\langle \Phi_0 | \mathrm{e}^{\hat{J}} \, \hat{H} \, \mathrm{e}^{\hat{J}} | \Phi_0 \rangle}{\langle \Phi_0 | \mathrm{e}^{2\hat{J}} | \Phi_0 \rangle}, \quad |\Phi_T\rangle = \mathrm{e}^{\hat{J}} \, |\Phi_0\rangle$$

- 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**

Similarity Transformation – Transcorrelated (TC) Method

Describe the cusp condition and/or capture part of correlation with a correlated wavefunction Ansatz → incorporate into Hamiltonian!
Instead of $\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle$ solve the similarity transformed (ST) problem

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



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 $\hat{J}$!

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)

The Similarity Transformed TC Hamiltonian

Consequences:

- Sim. transf. $\bar{H}$ is non-Hermitian ($[\,\hat{H}, \hat{J}\,], \ldots$)
  $\rightarrow$ loss of variational principle
- 3-body interactions (and possibly higher order)

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

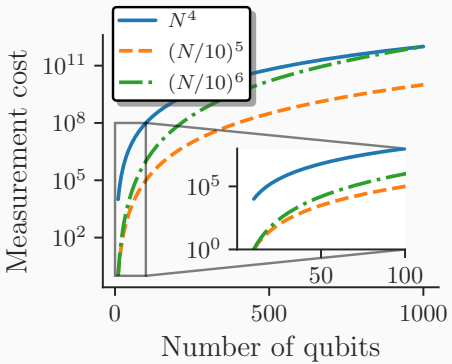
Rapid basis set convergence and more compact (right) eigenvector!

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

Scaling of TC – Measurement Cost

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

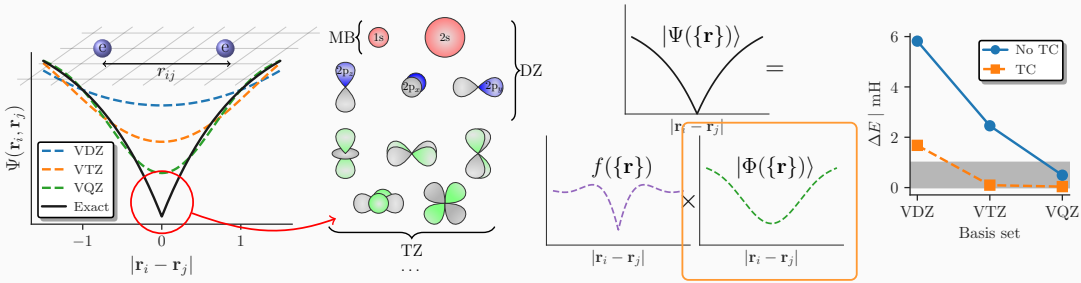
- Measurement formally scaling as N^6, with N being the number of orbitals
- Recently shown that N^6-scaling terms can be neglected to good accuracy*
- Current work on N^4-scaling approximation
- Order of magnitude less orbitals: since also no core functions needed in basis set†
- Shorter circuit depth, 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);

TC Motivation for quantum computing

Sharp cusp → necessitates **large basis set expansion**



Every spin-orbital needs a qubit → **Minimal basis** far from CBS results!

Transcorrelation (TC) enables **more accurate results with less spin-orbitals/qubits**

Intermezzo: Quantum Imaginary Time Evolution – QITE

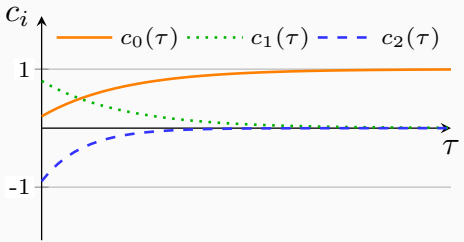
Since the TC Hamiltonian is non Hermitian, variational algorithms like VQE are not applicable!

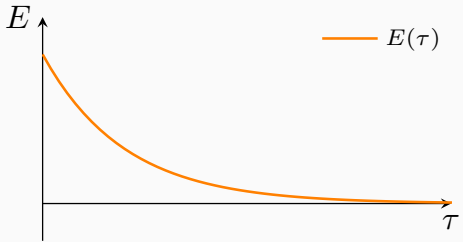
Quantum Imaginary Time Evolution – QITE

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

$$i\frac{\partial |\Psi\rangle}{\partial t} = \hat{H}|\Psi\rangle \quad \overset{\tau=it}{\rightarrow} \quad \frac{\partial |\Psi\rangle}{\partial \tau} = -\hat{H}|\Psi\rangle \quad \rightarrow \quad |\Psi(\tau)\rangle = N(\tau)\, \mathrm{e}^{-\hat{H}\tau}|\Psi(0)\rangle$$

$$|\Psi(0)\rangle = \sum_i c_i(0)|\psi_i\rangle \quad \rightarrow \quad |\Psi(\tau)\rangle = \mathrm{e}^{-\tau(\hat{H}-S_\tau)}\sum_i c_i(0)|\psi_i\rangle = \sum_i c_i(0)\,\mathrm{e}^{-\tau(E_i-S_\tau)}|\psi_i\rangle$$





Motta *et al.*, Nature Physics **16**, 205, 2020; *McArdle, *et al.*, npj Quantum Information **5**, 75, 2019;

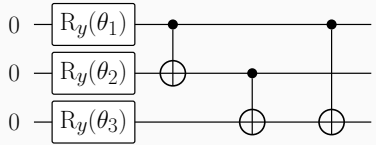
Ansatz-based QITE

$e^{-\hat{H}\tau}$ is **not unitary!** However, **Ansatz-based QITE**[*] allows to formulate non-unitary time evolution, as a minimization:

$$\frac{\partial \left|\Psi(\tau)\right\rangle}{\partial \tau} = -(\hat{H} - E(\tau)) \left|\Psi(\tau)\right\rangle, \quad \text{with} \quad E(\tau) = \langle\Psi(\tau)|\hat{H}|\Psi(\tau)\rangle,$$

(1) Use an Ansatz to approximate $|\Psi(\tau)\rangle$:

$$|\Psi(\tau)\rangle \approx |\Phi(\boldsymbol{\theta}(\tau))\rangle, \quad \text{with} \quad |\Phi(\boldsymbol{\theta}(\tau))\rangle = \hat{U}_n(\theta_n(\tau)) \cdots \hat{U}_1(\theta_1(\tau)) |\mathbf{0}\rangle = \hat{U}(\boldsymbol{\theta}(\tau)) |\mathbf{0}\rangle$$



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

[*] McArdle, *et al.*, npj Quantum Information **5**, 75, 2019; Motta *et al.*, Nature Physics **16**, 205, 2020; Sokolov, **WD**, Luo, Alavi, Tavernelli, arXiv:2201.03049; Zoufal *et al.*, arXiv:2108.00022 (2021);

Ansatz-based QITE

2) Use McLachlan's variational principle with Ansatz: $|\Psi(\tau)\rangle \approx |\Phi(\boldsymbol{\theta}(\tau))\rangle$

$$\delta || \frac{\partial}{\partial \tau} + \hat{H} - E(\tau) \, |\Phi(\boldsymbol{\theta}(\tau))\rangle \, || \stackrel{!}{=} 0$$

to express imaginary-time evolution of $|\Psi(\tau)\rangle$ as the change of the parameters $\frac{\partial \boldsymbol{\theta}(\tau)}{\partial \tau}$ wrt. to imaginary time τ:

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

with the **metric A** and **gradient C**:

$$A_{ij} = \frac{\partial \, \langle \Phi(\tau)|}{\partial \theta_i} \frac{\partial \, |\Phi(\tau)\rangle}{\partial \theta_j} \quad \text{and} \quad C_i = -\frac{\partial \, \langle \Phi(\tau)|}{\partial \theta_i} \hat{H} \, |\Phi(\tau)\rangle$$

*McArdle, *et al.*, npj Quantum Information **5**, 75, 2019; Motta *et al.*, Nature Physics **16**, 205, 2020; Sokolov, **WD**, Luo, Alavi, Tavernelli, arXiv:2201.03049; Zoufal *et al.*, arXiv:2108.00022 (2021);

Update in parameters

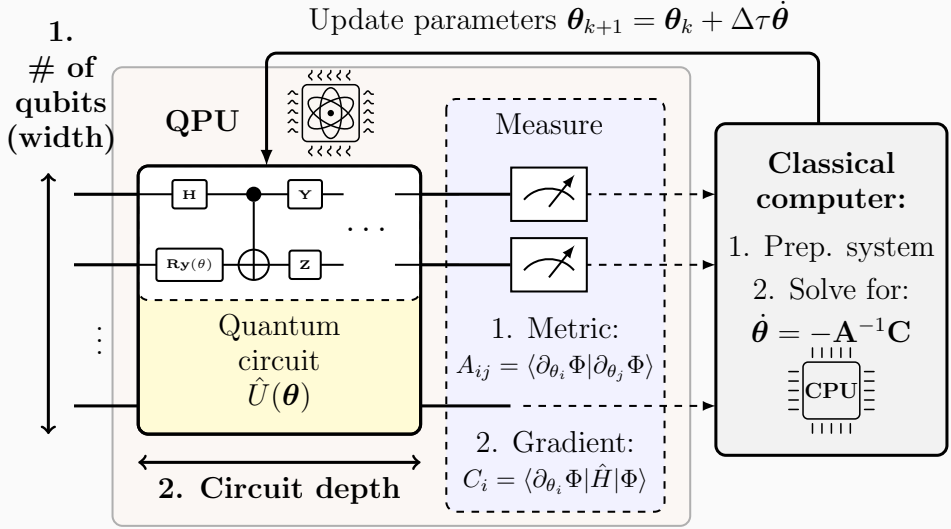
- No need for optimization → solution to linear system of equations
- Change in parameters with e.g. Euler method with timestep $\Delta\tau$

$$\boldsymbol{\theta}(\tau + \Delta\tau) \approx \boldsymbol{\theta}(\tau) + \dot{\boldsymbol{\theta}}\Delta\tau = \boldsymbol{\theta}(\tau) + \mathbf{A}^{-1}\boldsymbol{C}\Delta\tau$$

- Repeating this $N_\tau = \frac{\tau_{tot}}{\Delta\tau}$ times, allows to simulate imaginary time evolution for a total time τ_{tot}.
- To perform VarQITE **A** and **C** must be measured on a quantum computer → need to **measure derivatives** w.r.t. θ_i:
  Numerical differentiation, parameter-shift rule, linear combination of unitaries*

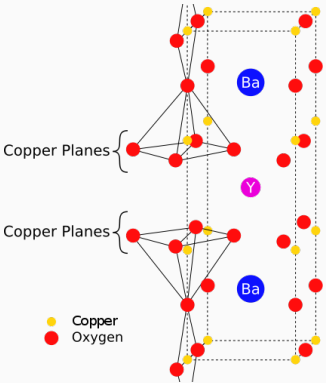
* Schuld *et al.*, Phys. Rev. A 99, 032331 (2019); Romero *et al.*, Quantum Science and Technology, 4, 1 (2019); Li and Benjamin, Phys. Rev. X 7, 021050 (2017);

QITE Workflow



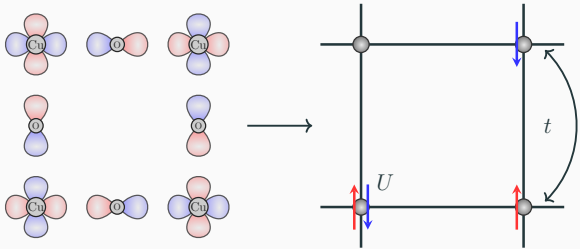
Applications: Circuit depth – Hubbard model

High-T_C Superconductors and the Hubbard Model



YBCO

Mapping to an effective lattice model:



The Hubbard Hamiltonian

$$\hat{H} = -t \sum_{\langle i,j \rangle, \sigma} \left(c^{\dagger}_{i,\sigma} c_{j,\sigma} + h.c. \right) + U \sum_{i} n_{i,\uparrow} n_{i,\downarrow}$$

Strong interaction $\Rightarrow$ highly multiconfigurational

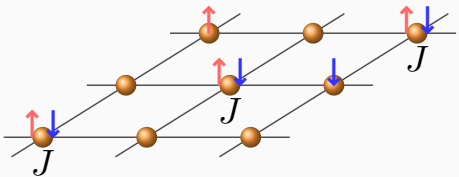
Hubbard, 1963; Kanamori, 1963; Gutzwiller; 1963; Anderson, 1987; Emery, 1987; Zhang and Rice, 1988; Bednorz and Müller, 1986

Similarity Transformation based on the Gutzwiller Ansatz

Reduce circuit depth with transcorrelated Ansatz

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

$$|\Psi\rangle = \mathrm{e}^{\hat{\tau}} |\Phi\rangle, \quad \hat{\tau} = J \sum_i n_{i\uparrow} n_{i\downarrow}$$



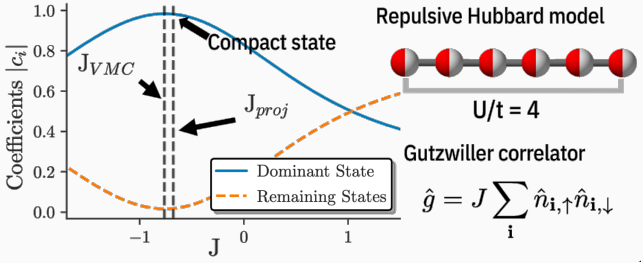
- Perform an exact *similarity transformation* (ST) of the Hubbard Hamiltonian $\hat{H}$:

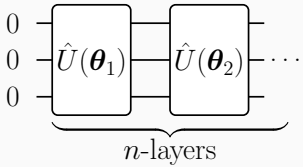
$$\mathrm{e}^{-\hat{\tau}} \hat{H} \mathrm{e}^{\hat{\tau}} = \bar{H} |\Phi\rangle = \left(-t \sum_{\langle i,j \rangle, \sigma} \mathrm{e}^{-\hat{\tau}} a^\dagger_{i\sigma} a_{j\sigma} \mathrm{e}^{\hat{\tau}} + U \sum_i n_{i\uparrow} n_{i\downarrow} \right) |\Phi\rangle = E |\Phi\rangle$$

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);

Increased compactness of right eigenvector

- Leads to a **non-Hermitian** operator with **3-body interactions** in a momentum space representation
- **Increased compactness** of the right EV, due to downfolding of correlations into Hamiltonian





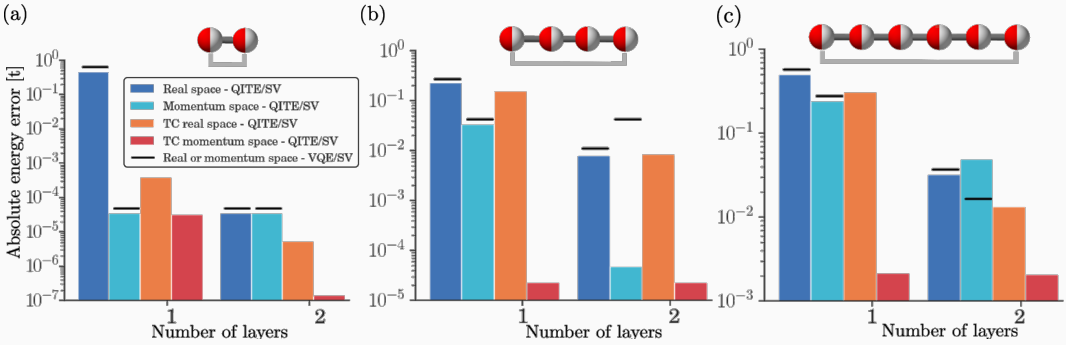
$\Rightarrow$ Does the increased compactness/more single reference character have an impact on the necessary Ansatz (depth)?

WD, Luo, Alavi, PRB, **99**, 075119 (2019); Sokolov, **WD**, Luo, Alavi, Tavernelli, arXiv:2201.03049

Results – Hubbard model

Increased compactness ⇒ less expressive Ansatz on quantum hardware necessary ⇒ **shorter quantum circuit**/less layers ⇒ cheaper and more resilient to noise

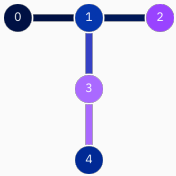
Statevector simulation – n-layer UCCSD Ansatz – $U/t = 4$

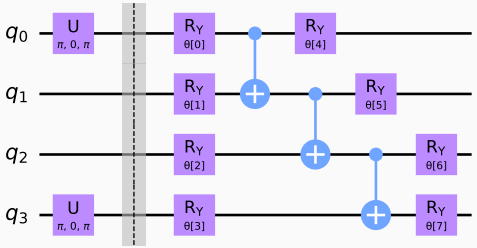


Black bars indicate VQE results

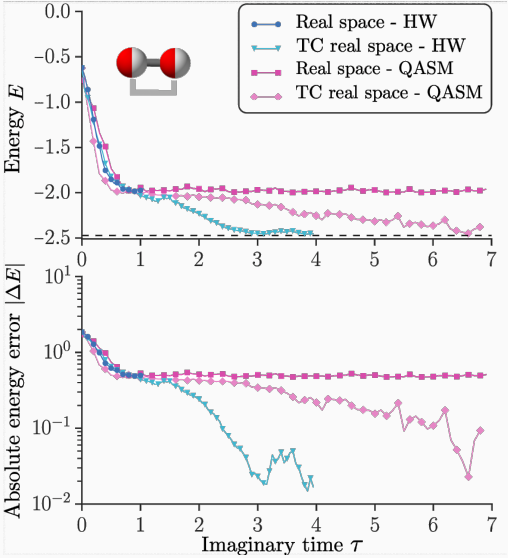
Actual experimental results for the Hubbard model on ibmq_lima

- 2-site Hubbard model





Hardware-efficient RY Ansatz



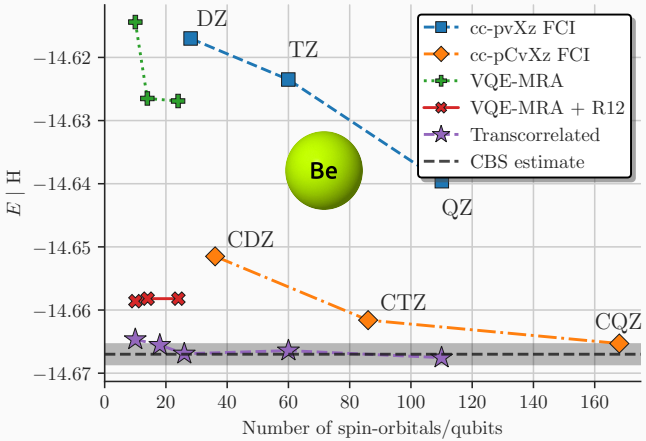
Applications: Circuit width – *Ab initio* problems

Can we reduce the number of necessary qubits (width)?

Beryllium atom

Beryllium atom – exact statevector simulations

VQE+MRA (+R12): (approx.) explicitly correlated method by Schleich *et al.*[*]



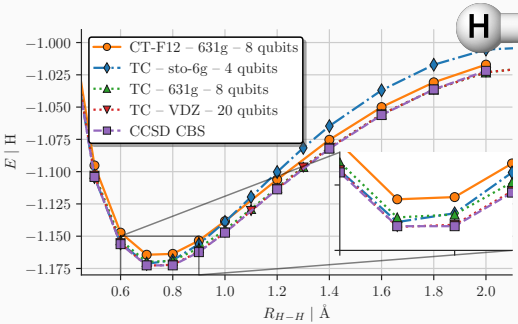
[*]VQE/MRA+[2]R12: Schleich *et al.*, Phys. Chem. Chem. Phys., 2022, 24, 13550 (2022)

Hydrogen molecule

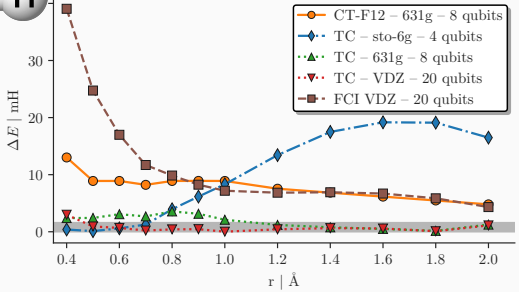
Favorite quantum chemistry test case: Hydrogen molecule – H_2

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

Exact statevector simulation – UCCSD Ansatz



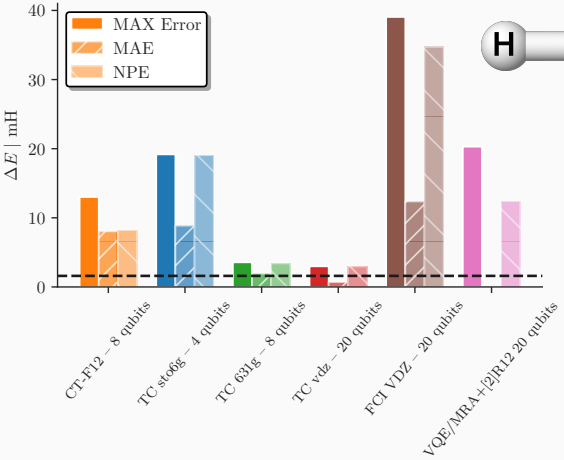
Energy vs. bond distance



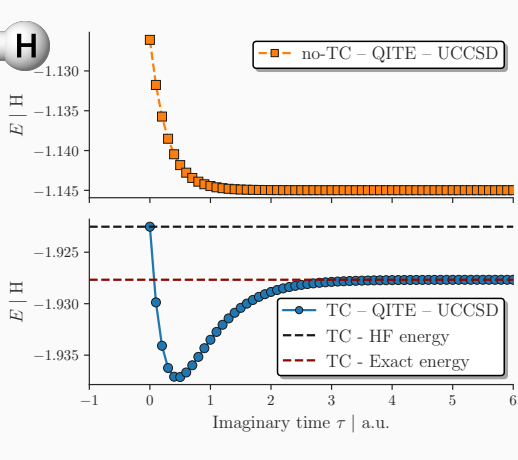
Error wrt. CBS result vs. bond distance

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

H_2 cont



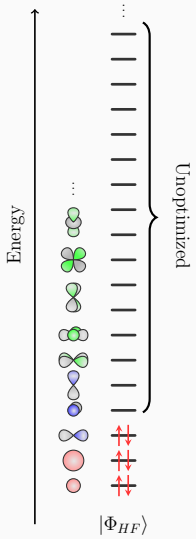
Error statistics



Imaginary time evolution – STO-6G – 0.7Å

CT-F12: Motta *et al.*, Phys. Chem. Chem. Phys. **22**, 24270, 2020 VQE/MRA+[2]R12: Schleich *et al.*, arXiv:2110.06812, 2021

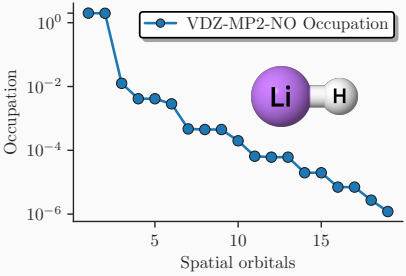
(Virtual) orbital optimization

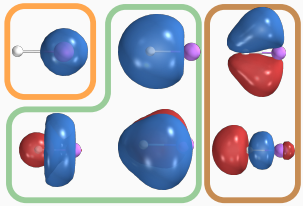


- "Standard basis sets" not optimized for the TC method

→ include effect of virtuals through orbital optimization / downfolding

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

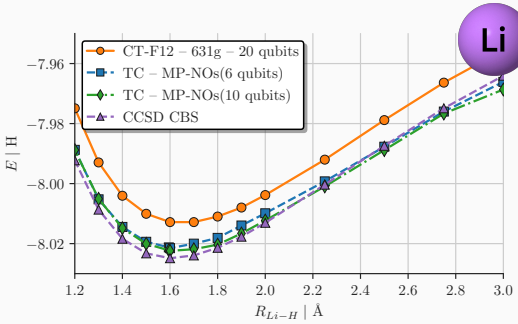




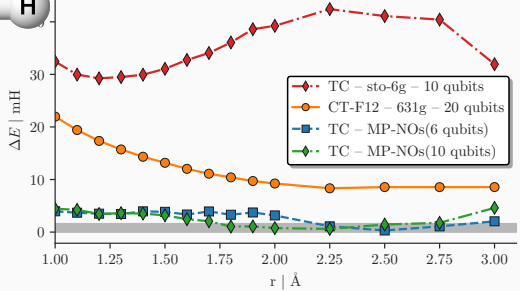
FNO: Sosa *et al.*, Chem. Phys. Lett. **159** (2-3) 148 (1989); Taube and Bartlett Collect. Czech. Chem. Commun. **70**, 837 (2005); **NISQ:** Gonthier *et al.* arXiv:2012.04001; (2020); Verma *et al.* JCP **155**, 034110 (2021)

LiH – Energy

Lithium hydride – exact statevector simulation – UCCSD Ansatz – Li 1s frozen



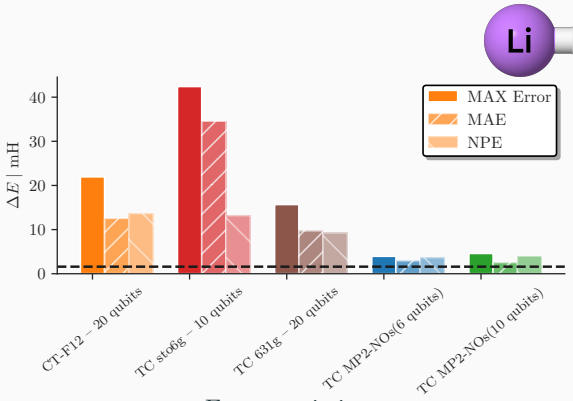
Energy vs bond distance



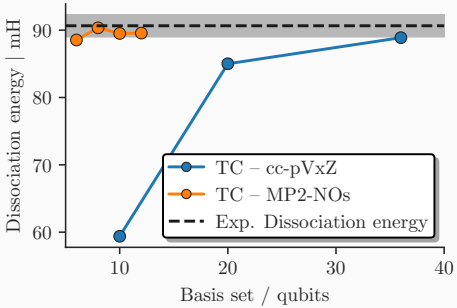
Error wrt. CBS result vs bond distance

LiH – Dissociation energy

Error statistics and comparison to **experimental*** dissociation energy



Error statistics

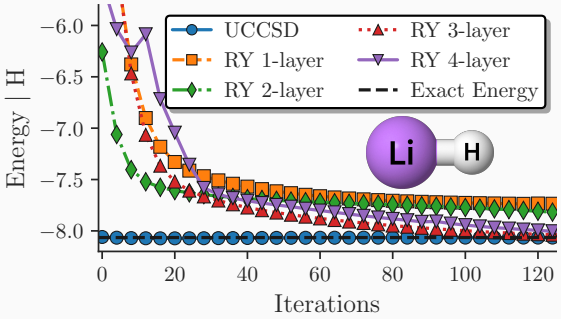


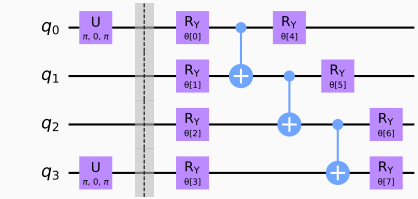
Experimental dissociation energy

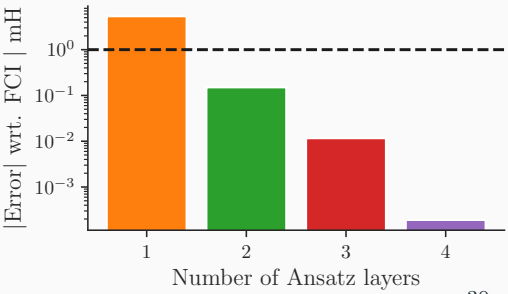
*Haeffler *et al.*, Phys. Rev. A, 1996, 53, 6, 4127 (1996)

LiH – Hardware-efficient Ansatz

- LiH at equilibrium bond distance with 3 MP2 NOs.
- Hardware efficient RY Ansatz with linear entangling layer and parity encoding.
- Statevector simulation



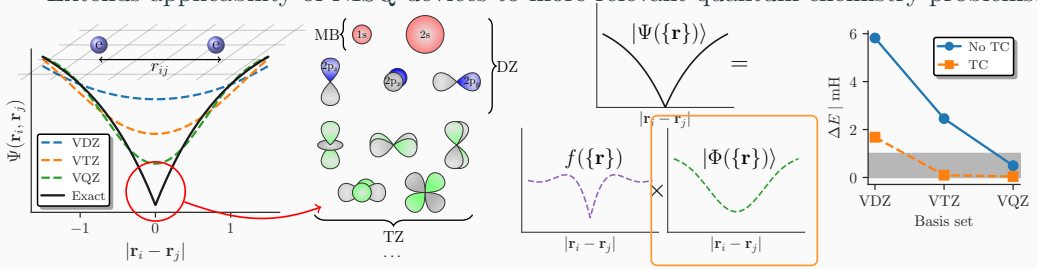




Conclusions and Outlook

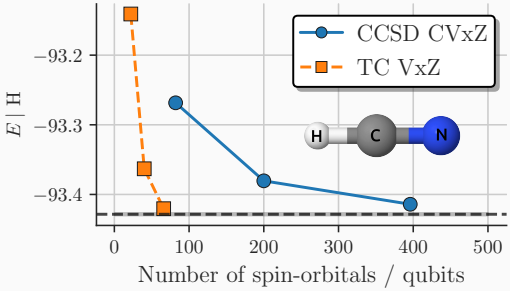
Conclusion – Transcorrelated Approach on Quantum Hardware

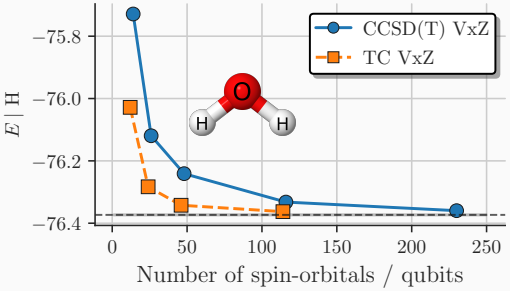
- The TC method partially transfers electronic correlations from the wavefunction into the Hamiltonian, exactly capturing the cusp condition.
- Non-Hermitian Hamiltonian requires quantum imaginary time evolution, instead of standard VQE.
- Reduce qubit requirements and circuit depth, due to accurate results with a small basis sets.
- Extends applicability of NISQ devices to more relevant quantum chemistry problems.



Outlook: Transcorrelated Approach on Quantum Hardware

Scaling to larger systems





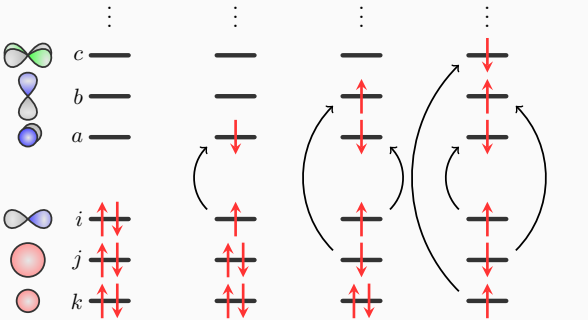
Open questions: **Noise**, approximations to metric **A**, which Ansatz to use? ...

Thank you for your attention!

Exponential scaling of Full Configuration Interaction

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

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

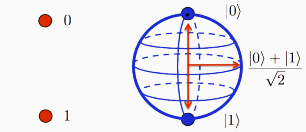


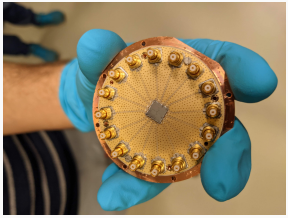
All possible excitations from HF determinant

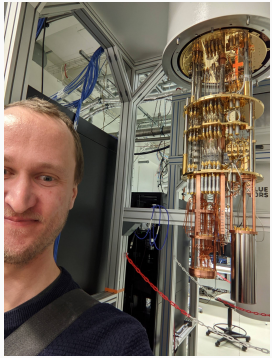
Number of possible states for given number of electrons, N, and orbitals, n, $\sim \binom{N}{n}$

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	$\sim 2 \cdot 10^9$

The case for quantum computing







@ Chalmers

n qubits can encode $\sim 2^n$ states. Quantum algorithms use **entanglement, superposition** and **interference** to find solution

$$\overbrace{(|0\rangle_1 + |1\rangle_1)}^{\text{qubit1}} \otimes \overbrace{(|0\rangle_2 + |1\rangle_2)}^{\text{qubit2}} = |00\rangle + |01\rangle + |10\rangle + |11\rangle$$

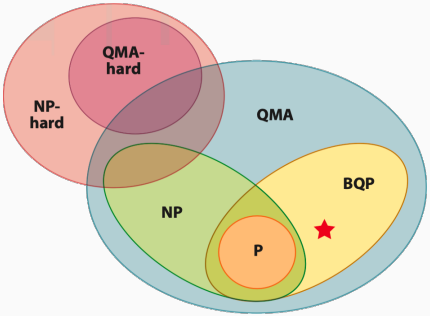
$$|q_1\rangle \otimes |q_2\rangle \otimes |q_3\rangle = \\ = |000\rangle + |001\rangle + |010\rangle + |100\rangle \\ + |011\rangle + |101\rangle + |110\rangle + |111\rangle$$

Benioff, J. Stat. Phys. **22** (5), 563 (1980); Feynman, Int. J. Theo. Phys. **21** (6/7), 467 (1982); Preskill, Quantum **2**, 79 (2018);

The case for quantum computing

Quantum chemistry potential use-case / killer-application of noisy intermediate-scale quantum (NISQ) devices and "quantum advantage" for relevant systems

- Efficient encoding of exponentially scaling wavefunction
- Effective measurement of Hamiltonian expectation values
- A system with > 60 qubits cannot be simulated with a classical computer
- A moderately-sized quantum processor (≈ 100 qubits) could outperform supercomputers for accurate solutions
- But how do we representand solve the problem on quantum hardware? **Ansätze** and **algorithms**



Kassal *et al.*, Annu. Rev. Phys. Chem. 2011. 62:185 (2010)

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 ≈ 80 orbitals

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

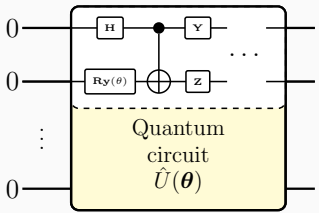
Variational Ansatz-based QITE – VarQITE

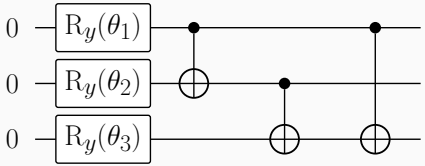
(Normalized) imaginary-time Schrödinger equation:

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

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

$$|\Psi(\tau)\rangle \approx |\Phi(\boldsymbol{\theta}(\tau))\rangle\,, \quad \text{with} \quad |\Phi(\boldsymbol{\theta}(\tau))\rangle = \hat{U}_n(\theta_n(\tau)) \cdots \hat{U}_1(\theta_1(\tau))\,|\mathbf{0}\rangle = \hat{U}(\boldsymbol{\theta}(\tau))\,|\mathbf{0}\rangle$$





State preparation $|\Psi(\boldsymbol{\theta})\rangle = \hat{U}(\boldsymbol{\theta})\,|\mathbf{0}\rangle$

McArdle *et al.*, npj Quantum, 5, 75 (2019); Yuan *et al.* Quantum 3, 191 (2019); Beach *et al.* PRB 100, 094434 (2019); Zoufal *et al.*, arXiv:2108.00022 (2021); McArdle *et al.* arXiv:2006.11181 (2019); Sokolov, **WD**, Luo, Alavi, Tavernelli, arXiv:2201.03049 (2022)

Three ingredients:

1. 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 $\delta\tau$: $\mathrm{e}^{-\delta\tau(\hat{H}-S_\tau)} \approx (1-\delta\tau(\hat{H}-S_\tau))$

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

2. 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)

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:

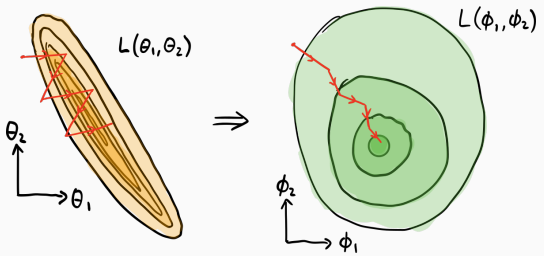
$$\sum_j A_{ij}\dot{\theta}_j = C_i, \quad \Longrightarrow \dot{\boldsymbol{\theta}} = \mathbf{A}^{-1}\boldsymbol{C}$$

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

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

$$\frac{dE(\tau)}{d\tau} = \langle\Phi(\tau)|\,\hat{H}\,\frac{d\,|\Phi(\tau)\rangle}{d\tau} = \sum_i \langle\Phi(\tau)|\,\hat{H}\,\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_i}\dot{\theta}_i \tag{6}$$
$$= -\sum_i C_i\dot{\theta}_i = -\sum_i C_i A_{ij}^{-1} C_j \leq 0,$$

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 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

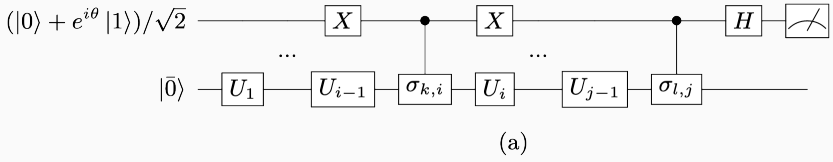
$$u_{1,i} = I \otimes \sigma_z, f_{1,i} = -i/4,$$
$$u_{2,i} = \sigma_z \otimes \sigma_z, f_{2,i} = i/4$$

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

Evaluation A and C with Quantum Circuits – cont.

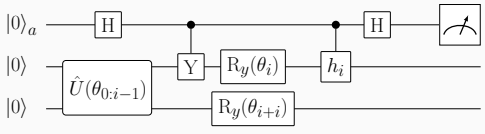
$$A_{ij} = \sum_{k,l} f^*_{k,i} f_{l,j} \langle \mathbf{0} | V'^{\dagger}_{k,i} V'_{l,j} | \mathbf{0} \rangle, \qquad C_i = \sum_{k,l} f^*_{k,l} \lambda_l \langle \mathbf{0} | V'^{\dagger} h_l \hat{V} | \mathbf{0} \rangle \tag{8}$$

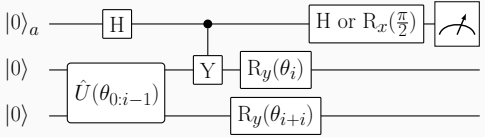
with $\hat{H} = \sum_l \lambda_l h_l$. Both $\mathbf{A}$ and $\boldsymbol{C}$ are of the form $a \cdot \mathrm{e}^{i\phi} \langle \mathbf{0} | \hat{U} | \mathbf{0} \rangle$ and can be evaluated on a quantum circuit.



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

Gradient in the Hermitian case: $C_i = \dfrac{\partial \langle \Phi |}{\theta_i} \hat{H} |\Phi\rangle$





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, arXiv:2201.03049 (2022); Schuld *et al.*, Phys. Rev. A 99, 032331 (2019)