Towards Real Chemical Accuracy on Current Quantum Hardware through the Transcorrelated Method

Werner Dobrautz

Chalmers University of Technology

Quo Vadis Electronic Structure Theory – QVEST
October 16, 2023

Outline

- Electronic Structure Theory – Quantum Chemistry
- The Case for Quantum Computing
- 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)
- Conclusions

Electronic Structure Theory – Quantum Chemistry

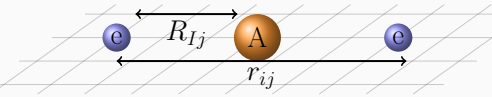
Ab Initio Quantum Chemistry – Electronic Structure Theory

To obtain insight on the **physical** and **chemical properties** (ground- and excited state energies, energy differences, response functions, ...) of quantum systems we need to **solve the Schrödinger equation**

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\,,$$

where all necessary information of a quantum system contained in electronic **molecular Hamiltonian**

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

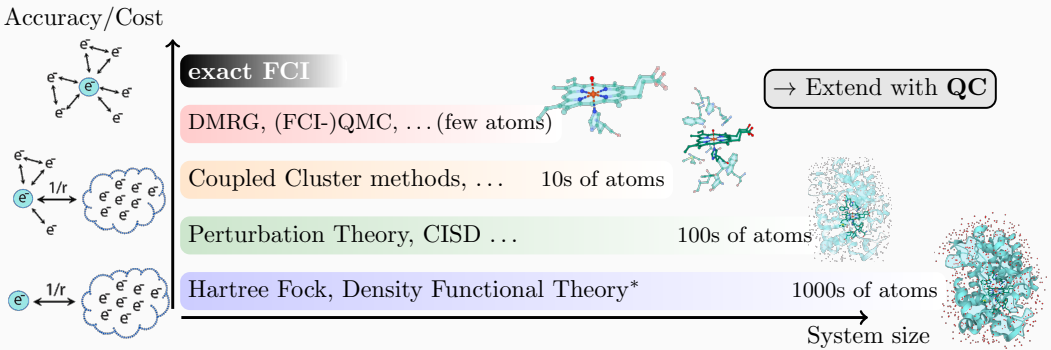


Coulomb repulsion correlates all electrons of a system → analytic solution too complex
→ **approximations and computational approaches**

Target: High accuracy to predict, interpret and compare with experimental results.

Accuracy and cost – scaling and hierarchy of methods

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



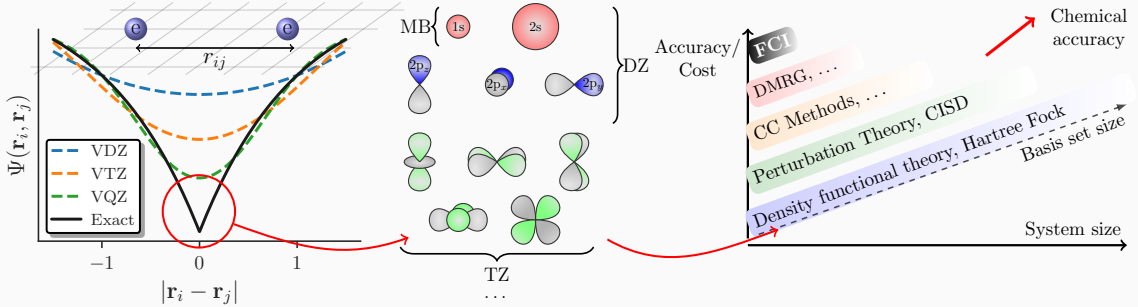
Need **highly-accurate** methods to describe **strongly correlated** problems

Cytochrome c: enzyme that eliminates toxic radicals produced by cells. From Santagati *et al.*, arXiv:2301.04114

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 ($r_{ij} = 0$)

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



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

The Case for Quantum Computing

Classical bit

0 1

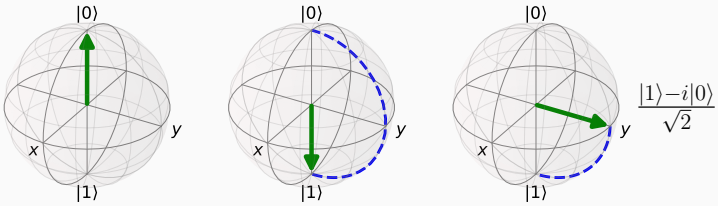
Quantum bit = qubit

$$a\,|0\rangle \quad + \quad b\,|1\rangle$$

Quantum bit = qubit

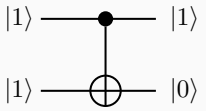
$$a\,|0\rangle \quad + \quad b\,|1\rangle$$
$$|a|^2 + |b|^2 = 1$$

Qubits – Bloch Sphere



Circuit model: Lines represent qubits and similar to classical circuits (AND, OR, ...) we can act with **operations/gates** on one (rotations) or **multiple qubits** (CNOT, ...)

Parametrized gates ($R_x(\theta)$, ...) and multi-qubit gates allow us to prepare entangled (non-classical) states, $|\Psi(\boldsymbol{\theta})\rangle = \hat{U}(\boldsymbol{\theta})\,|0\rangle$.



Multiple Qubits

Bringing **two** qubits together:

$$|\Psi\rangle = \overbrace{(|0\rangle + |1\rangle)}^{\text{qubit 1}} \otimes \overbrace{(|0\rangle + |1\rangle)}^{\text{qubit 2}} = |00\rangle + |01\rangle + |10\rangle + |11\rangle \qquad 4 \text{ states}$$

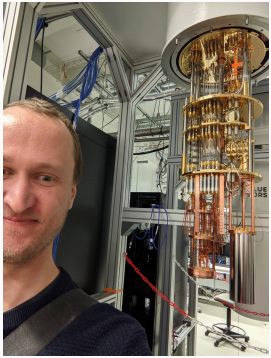
Three qubits:

$$\begin{aligned} |\Psi\rangle &= \overbrace{(|0\rangle + |1\rangle)}^{\text{qubit 1}} \otimes \overbrace{(|0\rangle + |1\rangle)}^{\text{qubit 2}} \otimes \overbrace{(|0\rangle + |1\rangle)}^{\text{qubit 3}} \\ &= |000\rangle + |001\rangle + |010\rangle + |100\rangle + |011\rangle + |101\rangle + |110\rangle + |111\rangle \qquad 8 \text{ states} \end{aligned}$$

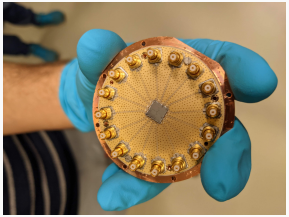
n qubits can encode exponentially many (2^n) states.

$\rightarrow$ Need new **quantum algorithms** to use this potential advantage!

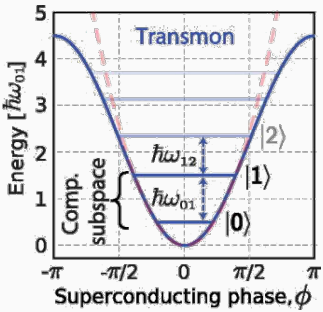
Noisy intermediate-scale quantum - NISQ



@ Chalmers



3 qubit device @Chalmers

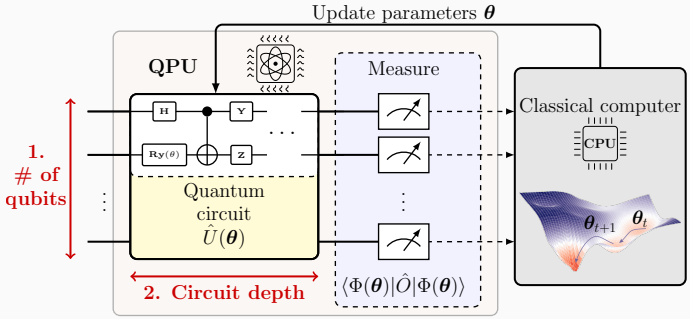


Qubits rely on quantum effects → very fragile, easily influenced by environmental effects/noise. Need to isolate and cool them close to absolute zero! Only few of them...

Kantz *et al.*. Applied Physics Reviews 6, 021318

Hybrid quantum-classical approach

Current quantum hardware has many flaws: **noise, decoherence and limited number of qubits** → **Hybrid quantum-classical approach**



Use pros of both CPUs and QPUs:

- Use **short-depth quantum circuits** that fit current hardware
- Can **improve on classical estimates** by non-classical states
- Store quantum state with **exponentially fewer resources**
- Use CPU to i.e. optimize gate parameters, $\boldsymbol{\theta}$

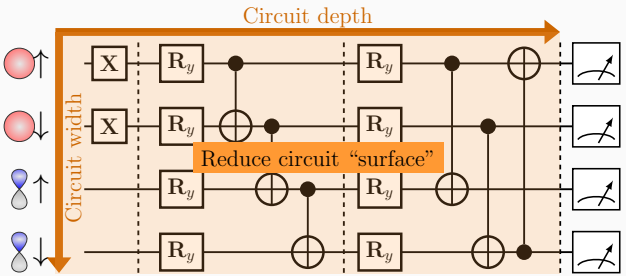
Limited number of qubits and circuit depth → small systems possible to study

Quantum Chemistry on Quantum Computers

1. Map fermionic Hamiltonian/basis functions onto quantum hardware/qubits*
2. Qubits encode occupation of spin-orbitals [0,1]
3. Use quantum algorithms for ground-, excited states, dynamics, ...

Prepare an initial state $|\Phi_0\rangle$

$$|\Phi_0\rangle = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$



Perform **unitary** operations of quantum algorithm on the qubits

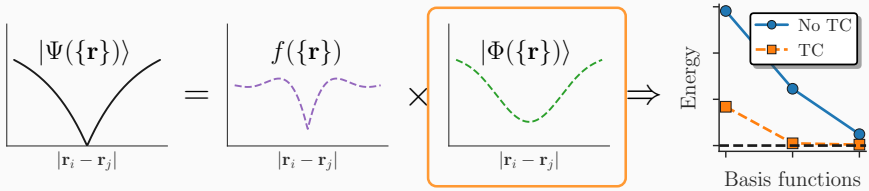
$$|\Phi\rangle = \hat{U}\,|\Phi_0\rangle = a_1 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \cdots + a_{2^N} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

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

Transcorrelation to reduce the computational footprint on quantum hardware

Cusp Condition – The Transcorrelated (TC) Method

Form of the **cusp** is known* → describe it with a **wavefunction Ansatz**



The transcorrelated (TC) method: use a Jastrow Ansatz, $e^{\hat{J}}$, with optimizable parameters J_{ij} (via VMC†) to transform the Hamiltonian:

$$|\Psi(\{\mathbf{r}\})\rangle = \exp\left[\sum_{ij} J_{ij} g(\tilde{r}_{ij})\right] |\Phi(\{\mathbf{r}\})\rangle \quad \rightarrow \quad \hat{H}\,|\Psi\rangle = E\,|\Psi\rangle \quad \rightarrow \quad \overbrace{e^{-\hat{J}}\,\hat{H}\,e^{\hat{J}}}^{\bar{H}_{\mathrm{TC}}}\,|\Phi\rangle = E\,|\Phi\rangle$$

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

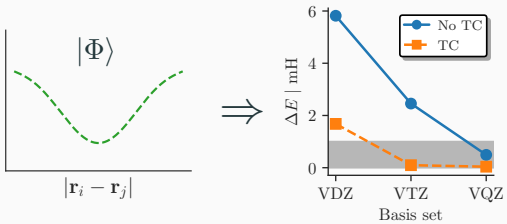
*Kato (1957); Boys and Handy (**1969**); Kutzelnigg (1985); **WD**, Luo, Alavi, PRB **99** (7), 075119 (2019); Cohen, Luo, Guther, **WD**, Tew, Alavi, JCP **151** (6), 061101 (2019); **WD**, Cohen, Alavi, Giner, JCP **156** (23), 234108 (2022); †Haupt, Hosseini, López Ríos, **WD**, Cohen and Alavi, JCP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

The Similarity Transformed TC Hamiltonian

Consequences:

- Transcorrelated $\bar{H}$ is not Hermitian! $\rightarrow$ loss of variational principle
- Additional 3-body terms in $\bar{H}$

Benefits: More accurate results with smaller basis sets/less qubits!

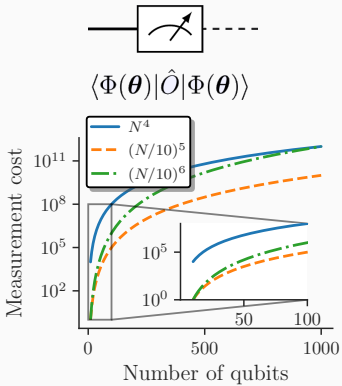


Cohen, Luo, Guther, **WD**, Tew, Alavi, JCP **151** (6), 061101 (2019); Haupt, Hosseini, López Ríos, **WD**, Cohen and Alavi, JCP **158**, 224105 (2023); Sokolov, **WD**, Luo, Alavi, Tavernelli, Physical Review Research 5 (2), 023174 (2023); **WD**, Sokolov, Liao, Lopez Rios, Rahm, Alavi, Tavernelli, arXiv:2303.02007 (2023)

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*
- **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!

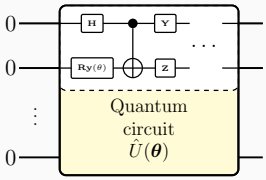
Variational Ansatz-based QITE – VarQITE

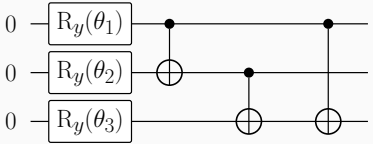
(Normalized) imaginary-time (Wick-rotated) Schrödinger equation

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

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:

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



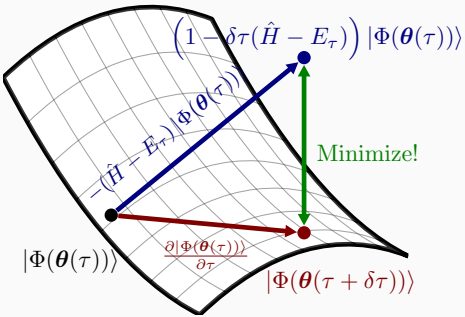


State preparation $|\Phi(\boldsymbol{\theta})\rangle = \hat{U}(\boldsymbol{\theta})\,|\mathbf{0}\rangle$

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

VarQITE

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



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

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

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

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

3. McLachlan's variational principle

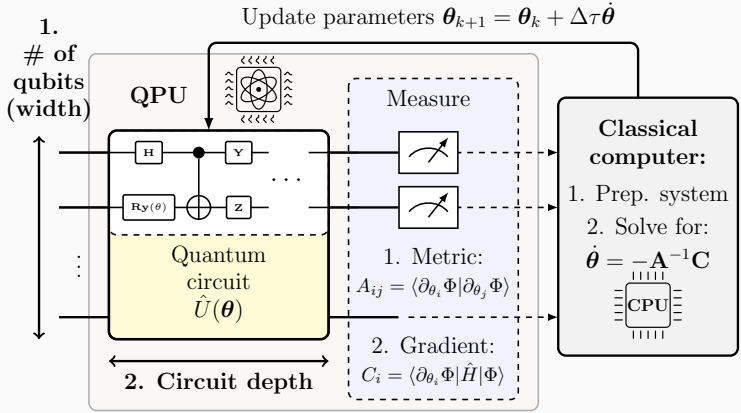
$$\delta\left|\left(\frac{\partial}{\partial\tau} + (\hat{H} - E_\tau)\right)|\Phi(\boldsymbol{\theta}(\tau))\rangle\right| = 0$$

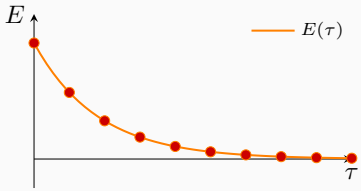
Evolution of parameters: $\dot{\boldsymbol{\theta}} = \mathbf{A}^{-1}\mathbf{C}, \quad A_{ij} = \frac{\partial\,\langle\Phi|}{\partial\theta_i}\frac{\partial\,|\Phi\rangle}{\partial\theta_j} \qquad C_i = -\frac{\partial\,\langle\Phi|}{\partial\theta_i}\hat{H}\,|\Phi\rangle$

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

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

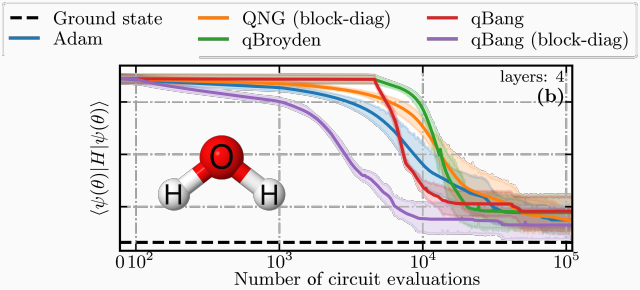
Improving classical optimization – qBANG – arXiv:2304.13882

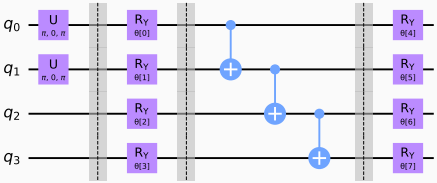
Measurement of Fisher information matrix, $\mathbf{A} \to \mathcal{O}(n_\theta^2)$ scaling!

Combines quasi-Newton approx. to the Hessian 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



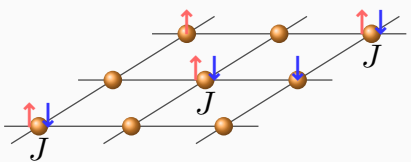


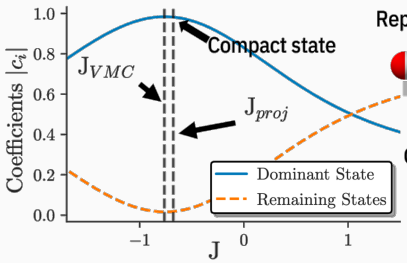
Results: Hubbard model and *ab initio* quantum chemistry problems

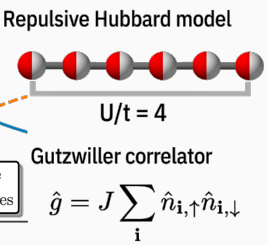
Reduce circuit depth with Transcorrelation

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

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







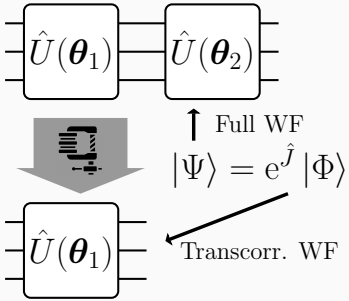
⇒ **Increased compactness** of the right EV, due to downfolding of correlations into Hamiltonian

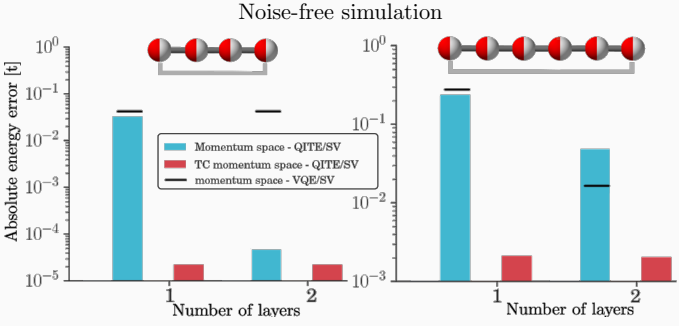
⇒ Does the increased compactness/more single reference character have an impact on the necessary quantum Ansatz depth?

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 circuit** necessary for accurate results!

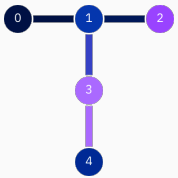


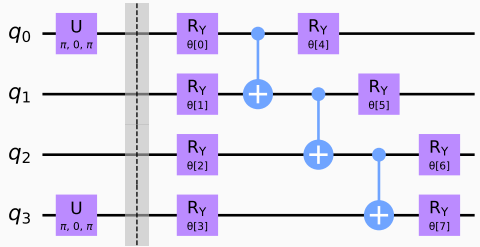


Noiseless statevector results, UCCSD Ansatz

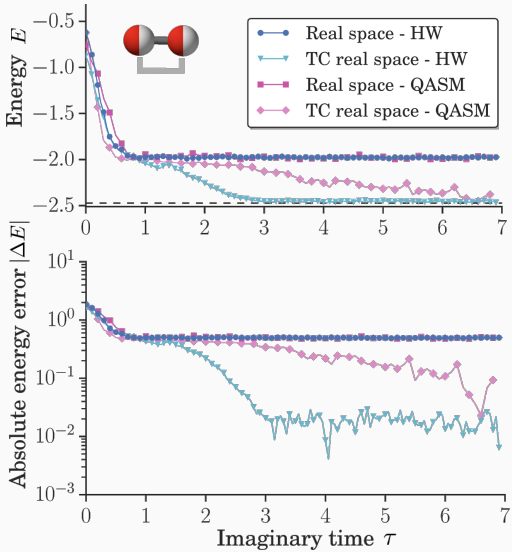
Actual experimental results for the Hubbard model on ibmq_lima

- 2-site Hubbard model
- Hardware-efficient Ansatz





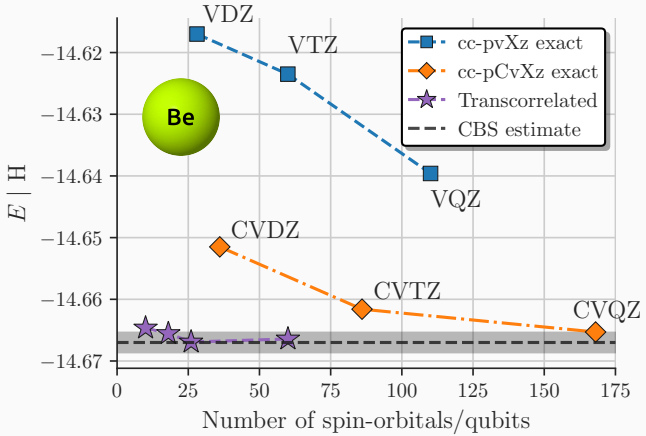
Hardware-efficient RY Ansatz



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

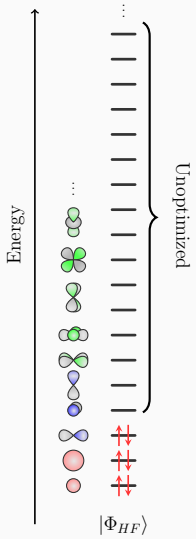
Beryllium atom – arXiv:2303.02007

Beryllium atom – exact simulation of a quantum device (no noise)
Goal: complete basis set (CBS) limit → full description to compare with experiment



WD, Sokolov, Liao, Lopez Rios, Rahm, Alavi, Tavernelli, arXiv:2303.02007 (2023)

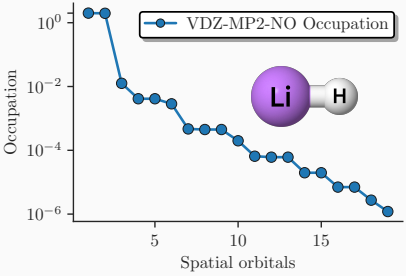
(Virtual) orbital optimization

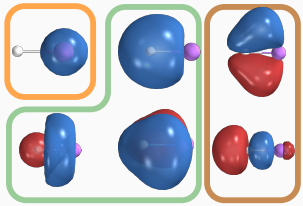


- "Standard basis sets" not optimized for the TC method

→ include effect of virtuals through orbital optimization / downfolding

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

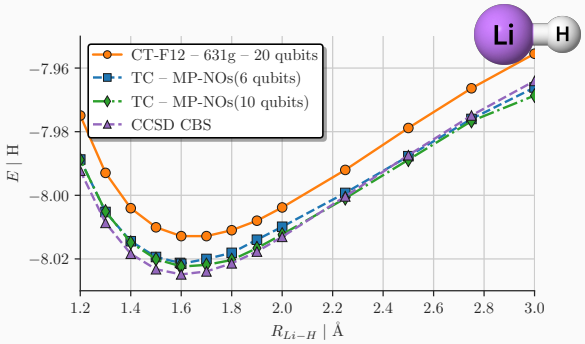




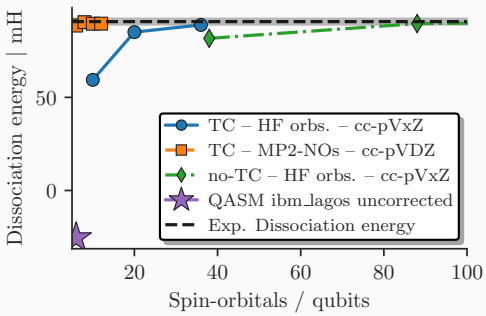
FNO: Sosa *et al.*, Chem. Phys. Lett. **159** (2-3) 148 (1989); Taube and Bartlett Collect. Czech. Chem. Commun. **70**, 837 (2005); **NISQ:** Gonthier *et al.* arXiv:2012.04001; (2020); Verma *et al.* JCP **155**, 034110 (2021)

Lithium hydride – LiH – arXiv:2303.02007

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, **arXiv:2303.02007** 2023

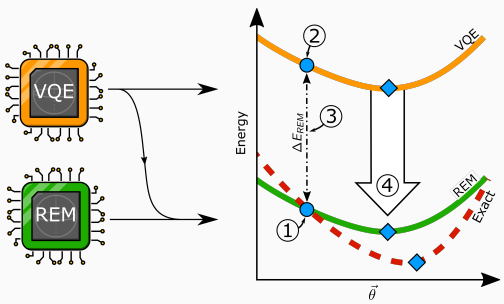
Reference-state Error Mitigation

JCTC Journal of Chemical Theory and Computation
pubs.acs.org/JCTC Article

Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry

Phalgun Lolur,[‖] Mårten Skogh,[‖] Werner Dobrautz, Christopher Warren, Janka Biznárová, Amr Osman, Giovanna Tancredi, Göran Wendin, Jonas Bylander, and Martin Rahm[*]

Cite This: *J. Chem. Theory Comput.* 2023, 19, 783−789

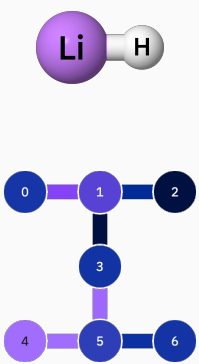


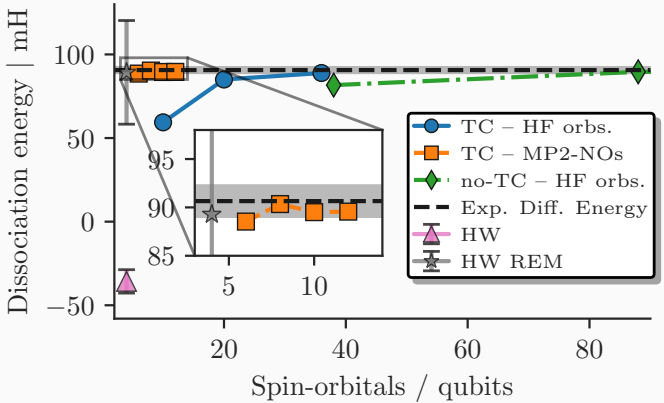
1. Classically compute exact reference energy (i.e. Hartree-Fock), $E(\boldsymbol{\theta}_{\text{ref}})$
2. Measure reference energy on noisy device $\mathcal{E}(\boldsymbol{\theta}_{\text{ref}})$, with reference parameters $\boldsymbol{\theta}_{\text{ref}}$
3. Calculate REM correction: $\Delta E_{\text{REM}}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

[*] 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 – arXiv:2303.02007

Hardware (HW) experiment: **lithium hydride** dissociation energy on `ibm_lagos`





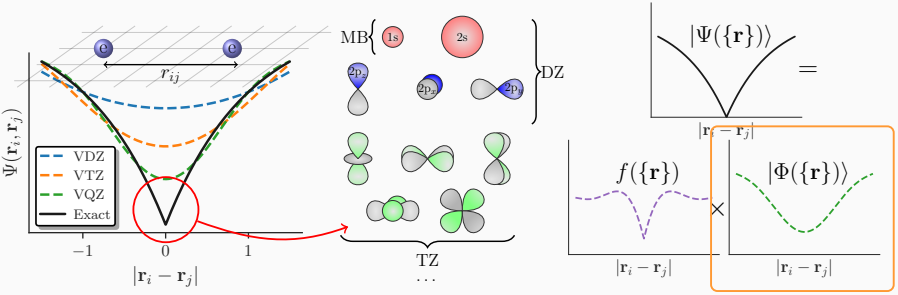
Hardware efficient RY Ansatz with linear entangling layer and parity encoding

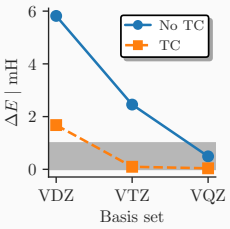
WD, Sokolov, Liao, Lopez Rios, Rahm, Alavi, Tavernelli, **arXiv:2303.02007** 2023; * Lolur, Skogh, **WD**, Warren, Biznárová, Osman, Tancredi, Wendin, Bylander, and Rahm, J. Chem. Theory Comput. 2023, **19**, 3, 783

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.
- With efficient error mitigation techniques → **extends applicability of current and near-term quantum devices** to more relevant quantum chemistry problems.





Acknowledgments

Ali Alavi

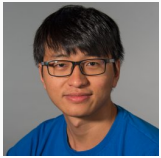
Igor O. Sokolov

Martin Rahm

Ivano Tavernelli

Pablo López Ríos

Hongjun Luo

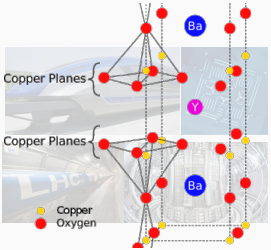


Ke Liao

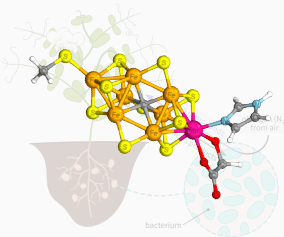
Funding:

Thank you for your attention!

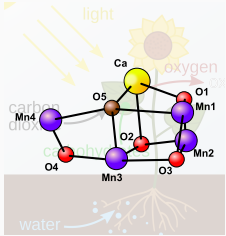
Applications of Quantum Algorithms: Electronic Structure Theory



YBCO: Unconventional high-T_c superconductivity



FeMoCo: primary cofactor of nitrogenase → nitrogen fixation



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

Surprisingly small systems responsible for interesting physical/chemical properties!

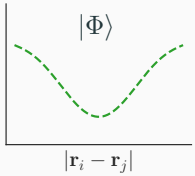
Strong electron correlation ⇒ challenging systems for computational approaches!

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

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:

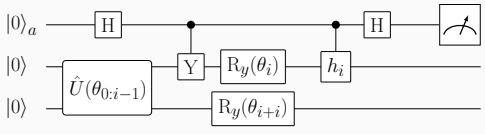
$$\bar{H} = \mathrm{e}^{-\hat{J}}\,\hat{H}\,\mathrm{e}^{\hat{J}} = \hat{H} + [\,\hat{H}, \hat{J}\,] + \frac{1}{2}\,[\,[\,\hat{H}, \hat{J}\,]\,, \hat{J}\,] + \ldots$$

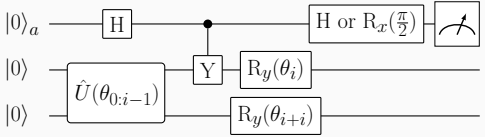
For the molecular Hamiltonian the BCH exp. terminates at 2nd order, as only kinetic energy operators in $\hat{H}$ do not commute with $\hat{J}$!

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)

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, Physical Review Research 5 (2), 023174 (2023) Schuld *et al.*, Phys. Rev. A 99, 032331 (2019)

VarQITE

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

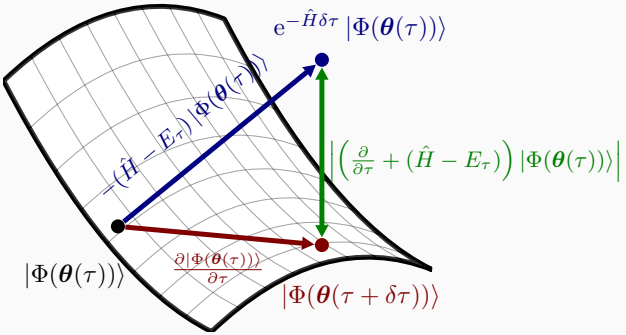
McLachlan, Molecular Physics, 8(1),39 (1964)

VarQITE

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

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

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



1. Ansatz $\left|\Phi(\boldsymbol{\theta}(\tau))\right\rangle = \hat{U}(\boldsymbol{\theta}(\tau))\left|0\right\rangle$
2. McLachlan's variational principle

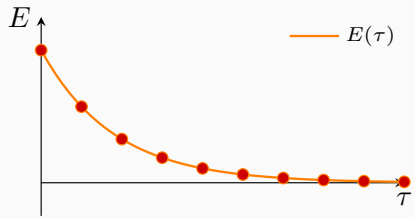
$$\delta\left|\left(\frac{\partial}{\partial\tau}+(\hat{H}-E_\tau)\right)\left|\Phi(\boldsymbol{\theta}(\tau))\right\rangle\right| = 0$$

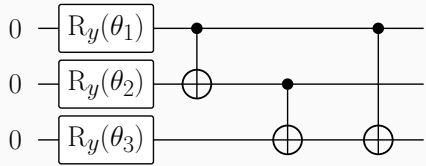
3. Evolution of parameters $\mathbf{A}\cdot\dot{\boldsymbol{\theta}} = C$

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

Quantum Imaginary Time Evolution

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





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

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:

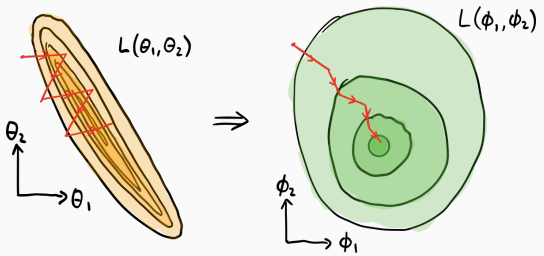
$$\sum_j A_{ij}\dot{\theta}_j = C_i, \quad \Longrightarrow \dot{\boldsymbol{\theta}} = \mathbf{A}^{-1}\boldsymbol{C}$$

with the metric:

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

and energy gradient:

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

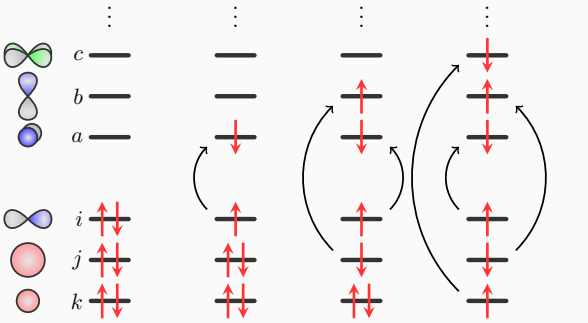


*Motta *et al.*, Nature Physics, 16, 205 (2020); pennylane.ai

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

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

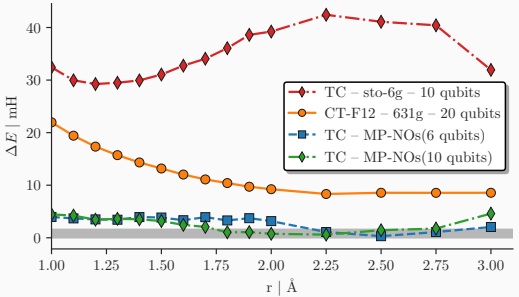
with

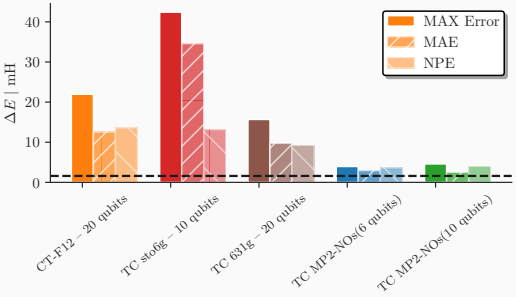
$$K_{rs}^{pq} = \langle \phi_p \phi_q | \hat{K} | \phi_r \phi_s \rangle$$
$$L_{stu}^{pqr} = \langle \phi_p \phi_q \phi_r | \hat{L} | \phi_s \phi_t \phi_u \rangle \quad \text{(48-fold symmetry in } L \text{ for real orbitals)}$$

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 ≈ 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)

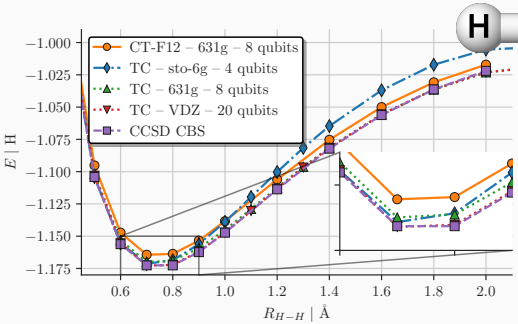

40
30
20
10
0
ΔE | mH
MAX Error
MAE
NPE
CT-F12 – 20 qubits
TC sto6g – 10 qubits
TC 631g – 20 qubits
TC MP2-NOs(6 qubits)
TC MP2-NOs(10 qubits)

Hydrogen molecule

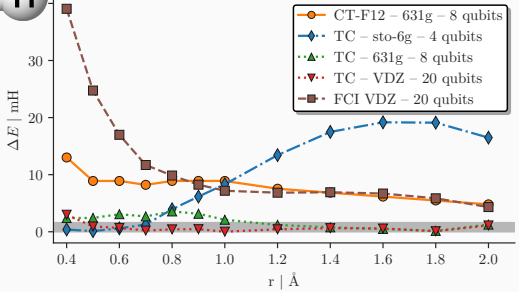
Favorite quantum chemistry test case: Hydrogen molecule – H_2

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

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

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.