What can DMRG learn from (post-)Hartree-Fock theory?

Sebastian Wouters

Workshop and Symposium on DMRG Technique
for Strongly Correlated Systems in Physics and Chemistry

Natal, June 27, 2015

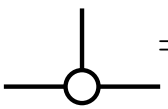
Contents

CHEMPS2: Spin-adapted DMRG for ab initio quantum chemistry

SW, W. Poelmans, P.W. Ayers and D. Van Neck, *Comput. Phys. Commun.* **185**, 1501 (2014)
SW and D. Van Neck, *EPJD* **68**, 272 (2014)

$$\mathbb{C}he[\mathcal{M}][\mathcal{P}][\mathcal{S}]\ 2$$

- Symmetry group $SU(2) \otimes U(1) \otimes P$ with $P \in \{C_1, C_i, C_2, C_s, D_2, C_{2v}, C_{2h}, D_{2h}\}$



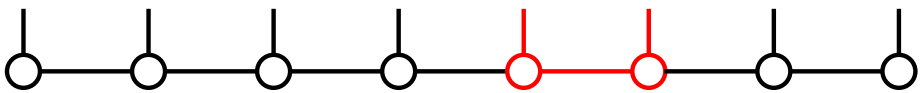
$$= A^{(ss^z NI)}_{(j_L j_L^z N_L I_L \alpha_L);(j_R j_R^z N_R I_R \alpha_R)} = \langle j_L j_L^z s s^z \mid j_R j_R^z \rangle \, \delta_{N_L+N,N_R} \delta_{I_L \otimes I, I_R} \, T^{(sNI)}_{(j_L N_L I_L \alpha_L);(j_R N_R I_R \alpha_R)} \quad (1)$$

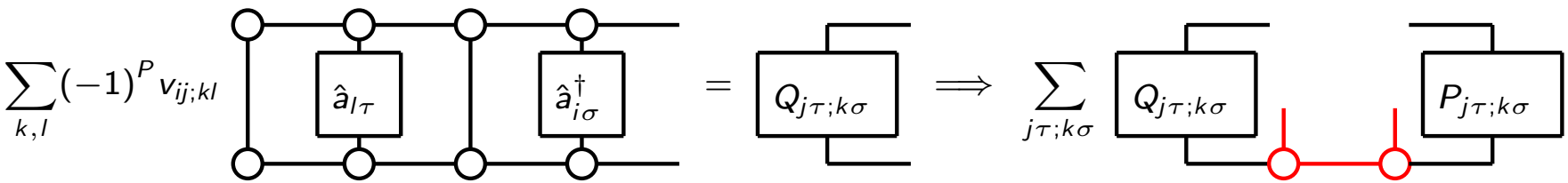
  - Block-sparsity whenever a Clebsch-Gordan coefficient is zero
  - Data compression for spin multiplets other than singlets
- c++ library with python interface
- GNU GPLv2
- https://github.com/sebwouters/chemps2
- Part of PSI4 and PYSCF (general purpose quantum chemistry packages)
- Hybrid MPI & OpenMP parallelization

CHEMPS2: Hybrid parallelization

- OpenMP over SU(2) $\otimes$ U(1) $\otimes$ P symmetry sectors in the virtual bonds
- MPI over (complementary) operator pairs

$$\hat{H} = \sum_{i,j,\sigma} t_{ij} \hat{a}^\dagger_{i\sigma} \hat{a}_{j\sigma} + \sum_{i,j,k,l,\sigma,\tau} v_{ij;kl} \hat{a}^\dagger_{i\sigma} \hat{a}^\dagger_{j\tau} \hat{a}_{l\tau} \hat{a}_{k\sigma} \tag{2}$$

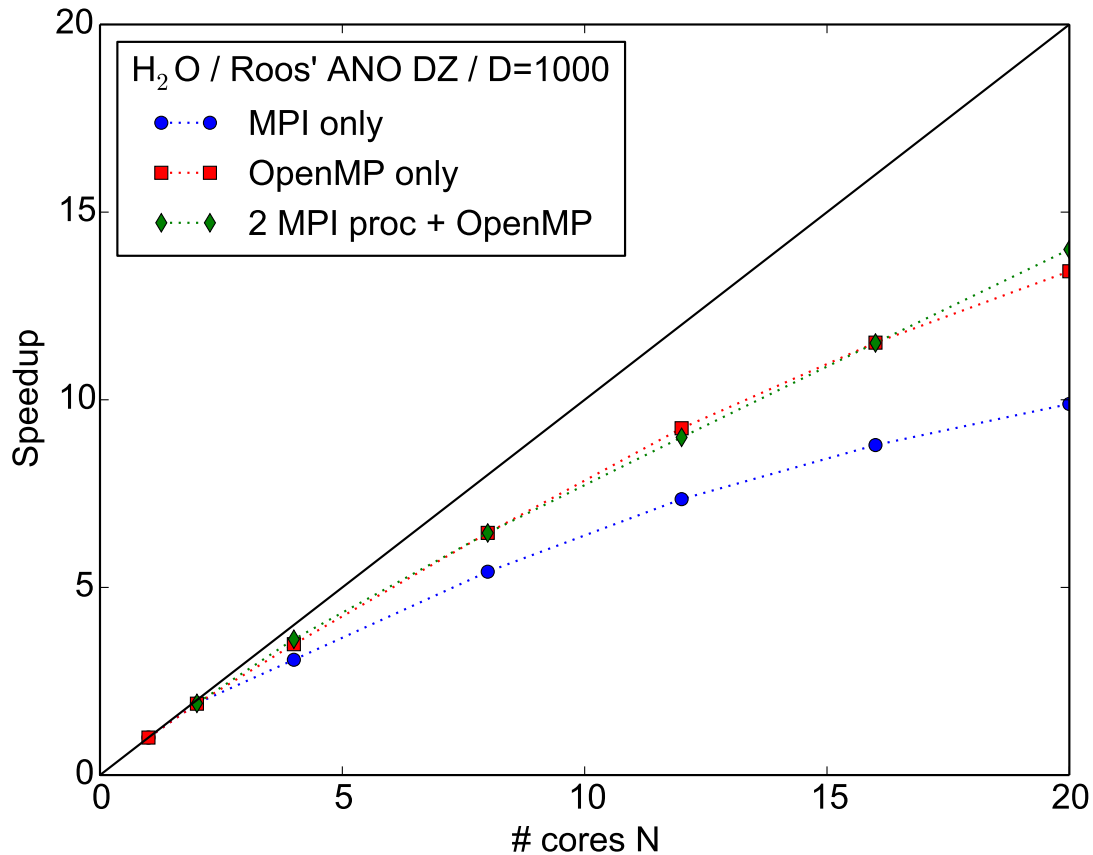




CHEMPS2: Scaling of the hybrid parallelization on 1 node

- H_2O in Roos' ANO DZ basis ($L = 41$ orbitals)
- Equilibrium: O @ (0, 0, 0) Å and H @ ($\pm$ 0.790689766, 0, 0.612217330) Å
- Intel Xeon Ivy Bridge (E5-2670 v2): 20 cores per node @ 2.5 GHz

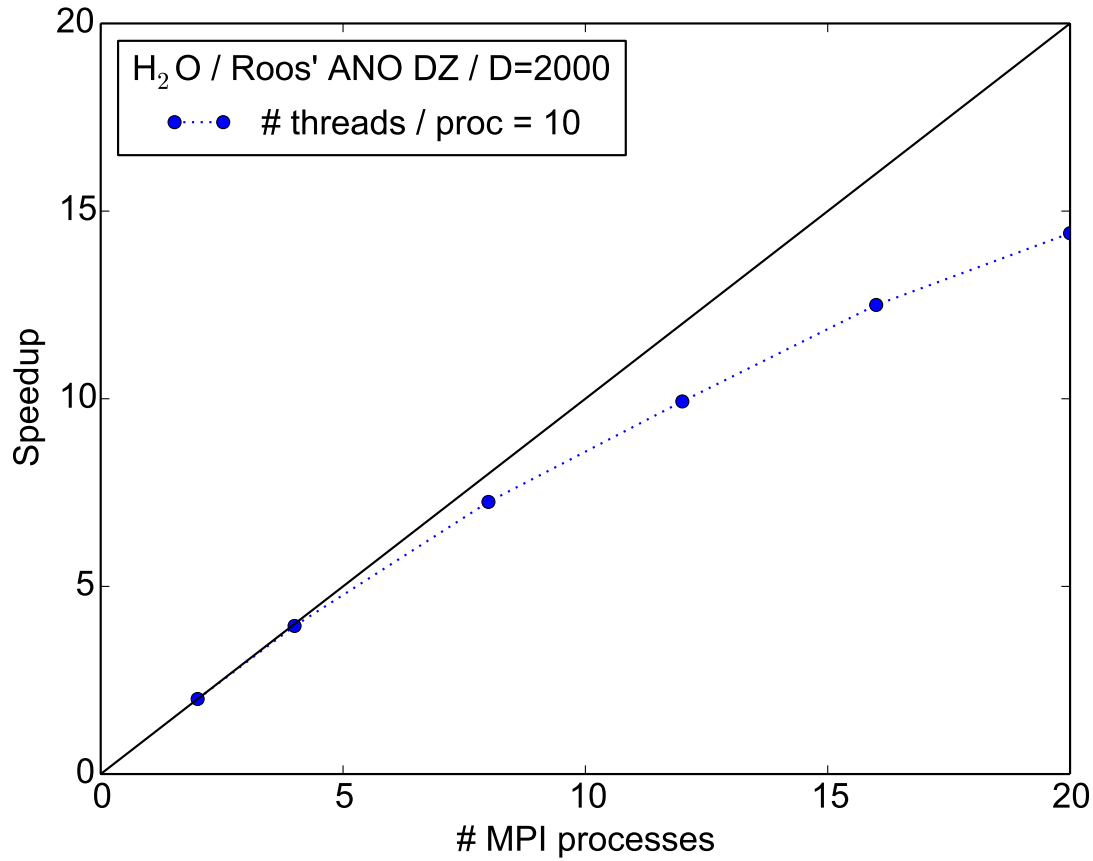
D	Wall time/sweep (s) [20 cores OpenMP]
1000	820
2000	4304
3000	11872
4000	23915



SW, work in progress (2015)

CHEMPS2: Scaling of the hybrid parallelization on multiple nodes

- 10 nodes, 2 sockets/node, 10 cores/socket
- Speedup of 100 on 200 cores



SW, work in progress (2015)

CHEMPS2: Capabilities

- Excited states by $\hat{H} = \hat{H}_0 + \sum_i \alpha_i \left|\Psi_i\right\rangle \left\langle\Psi_i\right|$
- Correlation functions and two-orbital mutual information

$$C_{ij}^{spin} = 4\left(\langle \hat{S}_i^z \hat{S}_j^z \rangle - \langle \hat{S}_i^z \rangle \langle \hat{S}_j^z \rangle\right) \quad (3)$$

$$C_{ij}^{spinflip} = \langle \hat{S}_i^+ \hat{S}_j^- \rangle + \langle \hat{S}_i^- \hat{S}_j^+ \rangle \quad (4)$$

$$C_{ij}^{dens} = \langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle \quad (5)$$

$$C_{ij}^{dirad} = \langle \hat{d}_{i\uparrow} \hat{d}_{j\downarrow} \rangle + \langle \hat{d}_{i\downarrow} \hat{d}_{j\uparrow} \rangle - \langle \hat{d}_{i\uparrow} \rangle \langle \hat{d}_{j\downarrow} \rangle - \langle \hat{d}_{i\downarrow} \rangle \langle \hat{d}_{j\uparrow} \rangle \quad (6)$$

$$\hat{d}_{i\sigma} = \hat{n}_{i\sigma}(1 - \hat{n}_{i\ -\sigma}) \quad (7)$$

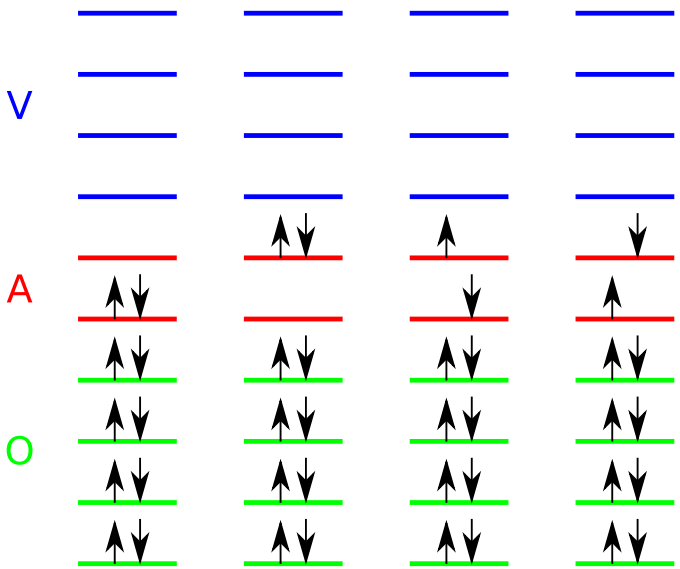
$$I_{ij} = \frac{1}{2}\left(S_1(i) + S_1(j) - S_2(ij)\right)\left(1 - \delta_{ij}\right) \quad (8)$$

- Reduced two-particle density matrix

$$\Gamma^A_{ij;kl} = \sum_{\sigma\tau} \langle \hat{a}^\dagger_{i\sigma} \hat{a}^\dagger_{j\tau} \hat{a}_{l\tau} \hat{a}_{k\sigma} \rangle \longrightarrow \text{DMRG-SCF} \quad (9)$$

$$\Gamma^B_{ij;kl} = \sum_{\sigma\tau} (-1)^{\sigma-\tau} \langle \hat{a}^\dagger_{i\sigma} \hat{a}^\dagger_{j\tau} \hat{a}_{l\tau} \hat{a}_{k\sigma} \rangle \quad (10)$$

DMRG-SCF



Static correlation:

- More than 1 important Slater determinant
- e.g. difluoride
  $|F_2\rangle \approx \alpha_0 |F^0 - F^0\rangle + \alpha_+ |F^+ - F^-\rangle + \alpha_- |F^- - F^+\rangle$
- Exact treatment

Dynamic correlation:

- Perturbation

DMRG-SCF:

- Occupied, Active, and Virtual orbital spaces
- A treats static correlation (by DMRG)
- OAV spaces optimized to minimize energy

DMRG-SCF algorithm

- Gradient and Hessian w.r.t. orbital rotations can be calculated with the 2-RDM Γ^A:

$$\hat{T}(\vec{x}) = \sum_{p<q} x_{pq} \sum_{\sigma} \left(\hat{a}^\dagger_{p\sigma} \hat{a}_{q\sigma} - \hat{a}^\dagger_{q\sigma} \hat{a}_{p\sigma} \right) \tag{11}$$

$$E(\vec{x}) = \langle e^{\hat{T}(\vec{x})} \hat{H} e^{-\hat{T}(\vec{x})} \rangle \approx E(0) + \vec{x}^T \vec{g} + \frac{1}{2} \vec{x}^T \mathbf{H} \vec{x} \tag{12}$$

- Newton-Raphson optimization of the orbitals $\left(\vec{x} = -\mathbf{H}^{-1}\vec{g}\right)$
- Hessian can have negative eigenvalues
  $\longrightarrow$ augmented Hessian Newton-Raphson method is more stable and used in practice

P.E.M. Siegbahn, J. Almlof, A. Heiberg and B.O. Roos, *JCP* **74**, 2384 (1981)
D. Zgid and M. Nooijen, *JCP* **128**, 144116 (2008)
D. Ghosh, J. Hachmann, T. Yanai and G.K.-L. Chan, *JCP* **128**, 144117 (2008)
SW, T. Bogaerts, P. Van Der Voort, V. Van Speybroeck and D. Van Neck, *JCP* **140**, 241103 (2014)

DMRG-SCF with CHEMPS2 is implemented

- As a standalone in CHEMPS2
- In PSI4: `https://github.com/psi4/psi4public`
- In PYSCF: `https://github.com/sunqm/pyscf`

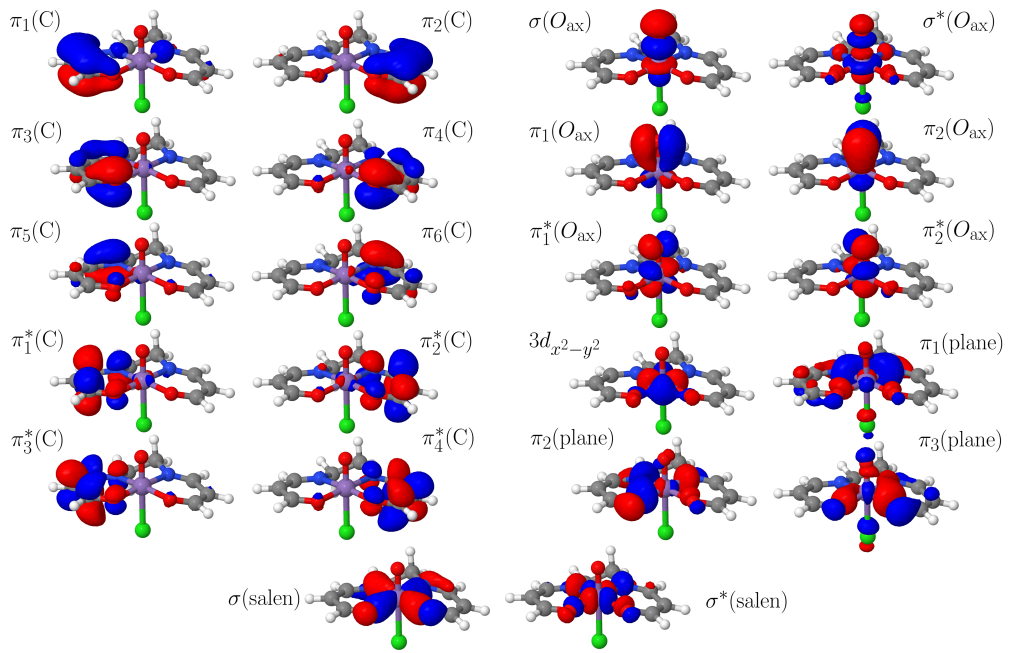
oxo-Mn(Salen)

SW, T. Bogaerts, P. Van Der Voort, V. Van Speybroeck and D. Van Neck, *JCP* **140**, 241103 (2014)

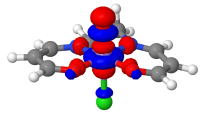
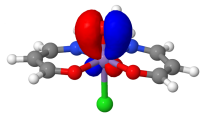
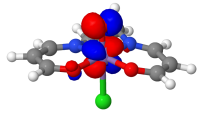
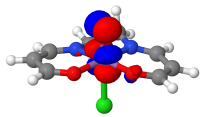
- Relative stability DMRG-SCF(28e, 22o) with D=4800 (extrapolated):

[kcal/mol]	6-31G*	cc-pVDZ	ANO-RCC-VDZP + DKH2
E(^3A) - E(^1A)	-5.0	-5.3	-4.0
E(^5A) - E(^1A)	14.5	12.1	14.5

- CASSCF(12e, 11o)/6-31G* : E(^5A) - E(^1A) = 42.9 kcal/mol ?
  J.S. Sears and C.D. Sherrill, *JCP* **124**, 144314 (2006)
- Dynamic correlation estimated to shift relative energies by $\approx$ 5 kcal/mol
  D. Ma, G.L. Manni and L. Gagliardi, *JCP* **135**, 044128 (2011)



oxo-Mn(Salen): Bond order of the axial oxygen atom

NOON		DMRG-SCF(28e, 22o)			CASSCF(12e, 11o)
		^1A	^3A	^5A	^5A
$\sigma(O_{ax})$		↑↓ (1.91)	↑↓ (1.90)	↑↓ (1.89)	↑↓
$\sigma^*(O_{ax})$		– (0.11)	– (0.11)	– (0.12)	–
$\pi_1(O_{ax})$		↑↓ (1.86)	↑↓ (1.77)	↑↓ (1.94)	↑↓
$\pi_2(O_{ax})$		↑↓ (1.85)	↑↓ (1.95)	↑↓ (1.94)	↑
$\pi_1^*(O_{ax})$		– (0.17)	↑ (1.04)	↑ (1.05)	↑
$\pi_2^*(O_{ax})$		– (0.17)	– (0.24)	↑ (1.04)	↑
$3d_{x^2-y^2}$		↑↓ (1.97)	↑ (1.00)	↑ (1.00)	↑
$\pi(C)$		↑↓	↑↓	↑ (1.01)	↑↓ (occ)
Bond order O_{ax}		3	2.5	2	1.5

Contents

Thouless' theorem: For Hartree-Fock theory

$$|\Phi(X,\overline{X})\rangle = \exp\left(\sum_{o\in occ;v\in virt} X^{vo}\hat{a}_v^\dagger\hat{a}_o - \overline{X}^{vo}\hat{a}_o^\dagger\hat{a}_v\right)|\Phi^0\rangle \tag{13}$$

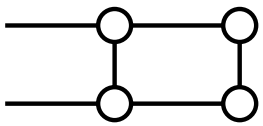
- A nonredundant parametrization of the entire Slater determinant manifold
- The (first order) tangent space is given by

$$|\Phi_{vo}\rangle = \frac{\partial}{\partial X^{vo}}|\Phi(X,\overline{X})\rangle\Big|_{X=0} = \hat{a}_v^\dagger\hat{a}_o|\Phi^0\rangle \tag{14}$$

- It reveals the elementary excitations of Hartree-Fock theory: particle-hole excitations
- Brillouin's theorem: $\langle\Phi^0 \mid \hat{H} \mid \Phi_{vo}\rangle = 0$
- Orthonormality: $\langle\Phi^0 \mid \Phi_{vo}\rangle = 0$ and $\langle\Phi_{vo} \mid \Phi_{wp}\rangle = \delta_{vw}\delta_{op}$
- Variational optimization in tangent space $\equiv$ Tamm-Dancoff approximation (TDA)
- Second order tangent space leads to configuration interaction method with singles and doubles (CISD)

Thouless' theorem: For MPS (1)

J. Haegeman, J.I. Cirac, T.J. Osborne, I. Pizorn, H. Verschelde and F. Verstraete, *PRL* **107**, 070601 (2011)

- Make all MPS tensors $[\mathbf{A}_0[i]]_{(\alpha_{i-1}n_{i\uparrow}n_{i\downarrow});\alpha_i}$ of the reference left-unitary
- Consider $\left[\mathbf{A}_0[i]\tilde{\mathbf{Q}}[i]\right]$ unitary with $\tilde{\mathbf{Q}}[i]$ the discarded renormalized basis states
- $\mathbf{r}[i] =$  (right environment of $\mathbf{A}_0[i]$)
- $\mathbf{B}(\mathbf{x})[i] = \tilde{\mathbf{Q}}[i]\mathbf{x}[i]\mathbf{r}[i]^{-\frac{1}{2}}$

SW, N. Nakatani, D. Van Neck and G.K.-L. Chan, *PRB* **88**, 075122 (2013)

- Grassmann manifold parametrization:

$$\mathbf{A}(\mathbf{x},\overline{\mathbf{x}})[i] = \exp\left(\mathbf{B}(\mathbf{x})[i]\mathbf{A}_0[i]^\dagger - \mathbf{A}_0[i]\mathbf{B}(\mathbf{x})[i]^\dagger\right)\mathbf{A}_0[i] \tag{15}$$

$$|\Phi(X,\overline{X})\rangle = \exp\left(\sum_{o\in occ;v\in virt} X^{vo}\hat{a}_v^\dagger\hat{a}_o - \overline{X}^{vo}\hat{a}_o^\dagger\hat{a}_v\right)|\Phi^0\rangle \tag{13}$$

- Nonredundant parametrization of the MPS manifold (same D_i as $\mathbf{A}_0[i]$):

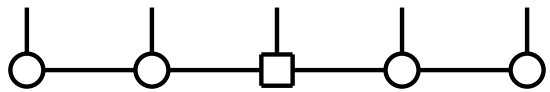
$$\mathbf{A}(\mathbf{x},\overline{\mathbf{x}})[1]^{n_{1\uparrow}n_{1\downarrow}}\mathbf{A}(\mathbf{x},\overline{\mathbf{x}})[2]^{n_{2\uparrow}n_{2\downarrow}}\ldots\mathbf{A}(\mathbf{x},\overline{\mathbf{x}})[L]^{n_{L\uparrow}n_{L\downarrow}} \tag{16}$$

Thouless' theorem: For MPS (2)

$$|\Psi(\mathbf{x},\bar{\mathbf{x}})\rangle \;\Rightarrow\; \mathbf{A}(\mathbf{x},\bar{\mathbf{x}})[1]^{n_{1\uparrow}n_{1\downarrow}}\mathbf{A}(\mathbf{x},\bar{\mathbf{x}})[2]^{n_{2\uparrow}n_{2\downarrow}}\ldots\mathbf{A}(\mathbf{x},\bar{\mathbf{x}})[L]^{n_{L\uparrow}n_{L\downarrow}} \tag{16}$$

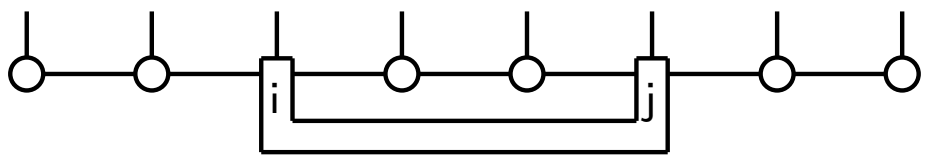
- Instead of optimizing over $\mathbf{A}[i]^{n_{1\uparrow}n_{1\downarrow}}$, we might as well optimize over $\mathbf{x}$
- The (first order) tangent space is given by

$$|\Psi_k\rangle = \frac{\partial}{\partial \mathbf{x}_k}|\Psi(\mathbf{x},\bar{\mathbf{x}})\rangle\Big|_{\mathbf{x}=\mathbf{0}} = \tag{17}$$



- It reveals the elementary excitations of DMRG: single-site excitations in which kept renormalized basis states are replaced with discarded ones
- Brillouin's theorem: $\langle\Psi^0 \mid \hat{H} \mid \Psi_k\rangle = 0$
- Orthonormality: $\langle\Psi^0 \mid \Psi_k\rangle = 0$ and $\langle\Psi_k \mid \Psi_l\rangle = \delta_{kl}$
- A variational optimization in the tangent space $\equiv$ DMRG-TDA
  B. Pirvu, J. Haegeman and F. Verstraete, *PRB* **85**, 035130 (2012)
- Second order tangent space leads to DMRG-CISD
  L. Vanderstraeten, J. Haegeman, T.J. Osborne and F. Verstraete, *PRL* **112**, 257202 (2014)

$$|\text{DMRG-CISD}\rangle = \sum_{i<j}$$



Thouless' theorem: DMRG-TDA example

N. Nakatani, **SW**, D. Van Neck and G.K.-L. Chan, *JCP* **140**, 024108 (2014)

- H_2O; aug-cc-pVDZ; R(OH) = 1.8111 a_0; $\angle$HOH = 104.45°; 1s frozen core
- DMRG-TDA is not good for excitations which involve multiple far-lying orbitals in the MPS chain
- HF-TDA fails for multi-particle excitations

	E_{conv}/E_h	$(E-E_{\mathrm{conv}})/mE_h$		
State	DMRG-4SA D=2000	DMRG-4SA D=500	DMRG-TDA D=500	EOM-CCSD
XA_1	-76.274 423	1.29	0.31	5.86
$2A_1$	-75.909 074	2.17	11.51	2.84
$3A_1$	-75.839 232	2.46	4.47	2.63
$4A_1$	-75.766 827	2.27	9.44	2.37
$1A_2$	-75.931 824	1.91	0.43	1.96
$2A_2$	-75.826 508	2.02	0.54	3.41
$3A_2$	-75.788 484	2.02	0.49	1.29
$1B_1$	-75.997 383	2.04	0.49	2.66
$2B_1$	-75.863 101	2.02	1.47	2.10
$3B_1$	-75.833 279	2.13	0.54	2.30
$4B_1$	-75.770 624	3.14	11.36	2.14
$1B_2$	-75.844 352	2.06	0.42	1.97
$2B_2$	-75.762 108	2.00	17.53	3.89

Thouless' theorem: DMRG-CISD example

SW, N. Nakatani, D. Van Neck & G.K.-L. Chan, *PRB* **88**, 075122 (2013)

- DMRG-TDA only captures excitations due to Brillouin's theorem
- DMRG-CISD also allows to improve the ground state
- Examples for spin-singlet ground states of the one-dimensional Hubbard chain L=8

U	N	Exact	DMRG $D_{SU(2)}=3$	DMRG $D_{SU(2)}=9$	DMRG-CISD $D_{SU(2)}=3$
0.1	8	-9.319312	-9.067465	-9.301264	-9.315185
1	6	-7.790647	-7.532068	-7.780764	-7.785715
10	4	-5.187427	-5.083270	-5.186955	-5.187090
100	4	-4.805753	-4.736845	-4.805615	-4.805360

Contents

MPS-AFQMC: A particular flavour of projector Monte Carlo

S. Wouters, B. Verstichel, D. Van Neck and G.K.-L. Chan, *PRB* **90**, 045104 (2014)

This idea was based on AFQMC with Slater determinants: S. Zhang, J. Carlson, and J.E. Gubernatis, *PRB* **55**, 7464 (1997)

- Consider the Hermitian operator $\hat{K} = e^{-\delta\tau\hat{H}}$
- The ground state can be found by repeated application of $\hat{K}$:

$$|\Phi_{GS}\rangle \propto \lim_{n\to\infty} \left(\hat{K}\right)^n |\Phi^{(0)}\rangle = \lim_{n\to\infty} |\Phi^{(n)}\rangle \qquad \text{if} \qquad \langle\Phi_{GS} \mid \Phi^{(0)}\rangle \neq 0$$

- The wavefunction is represented by an ensemble of walkers

$$|\Phi^{(n)}\rangle = \sum_{\phi} |\phi\rangle$$

- Decompose $\hat{K}$ in a probability distribution function $P(\mathbf{x})$ and operators $\hat{B}(\mathbf{x})$

$$\hat{K} = \sum_{\mathbf{x}} P(\mathbf{x})\hat{B}(\mathbf{x})$$

- Action of $\hat{B}(\mathbf{x})$ should not increase the complexity of the walkers $|\phi\rangle$

MPS-AFQMC: Hubbard-Stratonovich transformation for spin systems

- Consider $\hat{H} = \frac{1}{2}\sum_{ij} J_{ij} \hat{\vec{S}}_i . \hat{\vec{S}}_j + \sum_i h_i \hat{S}_i^z$ with $J_{ij} = \sum_k V_{ik}\gamma_k V_{jk}$ symmetric
- Define $\hat{v}_k^w = \sum_i \hat{S}_i^w V_{ik}\sqrt{-\gamma_k}$ with $w \in \{x, y, z\}$
- $\hat{\mathbf{v}} = (\hat{v}_1^x, \hat{v}_1^y, \hat{v}_1^z, \hat{v}_2^x, ...)$
- Hubbard-Stratonovich transformation for imaginary time evolution:

$$\hat{H} = \sum_i h_i \hat{S}_i^z - \frac{\hat{\mathbf{v}}^2}{2} = \hat{H}_1 - \frac{\hat{\mathbf{v}}^2}{2} \tag{18}$$

$$e^{-\delta\tau\hat{H}} = \int d\mathbf{x} P(\mathbf{x}) \hat{B}(\mathbf{x}) + \mathcal{O}(\delta\tau^2) \tag{19}$$

$$\hat{B}(\mathbf{x}) = e^{-\delta\tau\hat{H}_1/2} e^{\sqrt{\delta\tau}\mathbf{x}.\hat{\mathbf{v}}} e^{-\delta\tau\hat{H}_1/2} \tag{20}$$

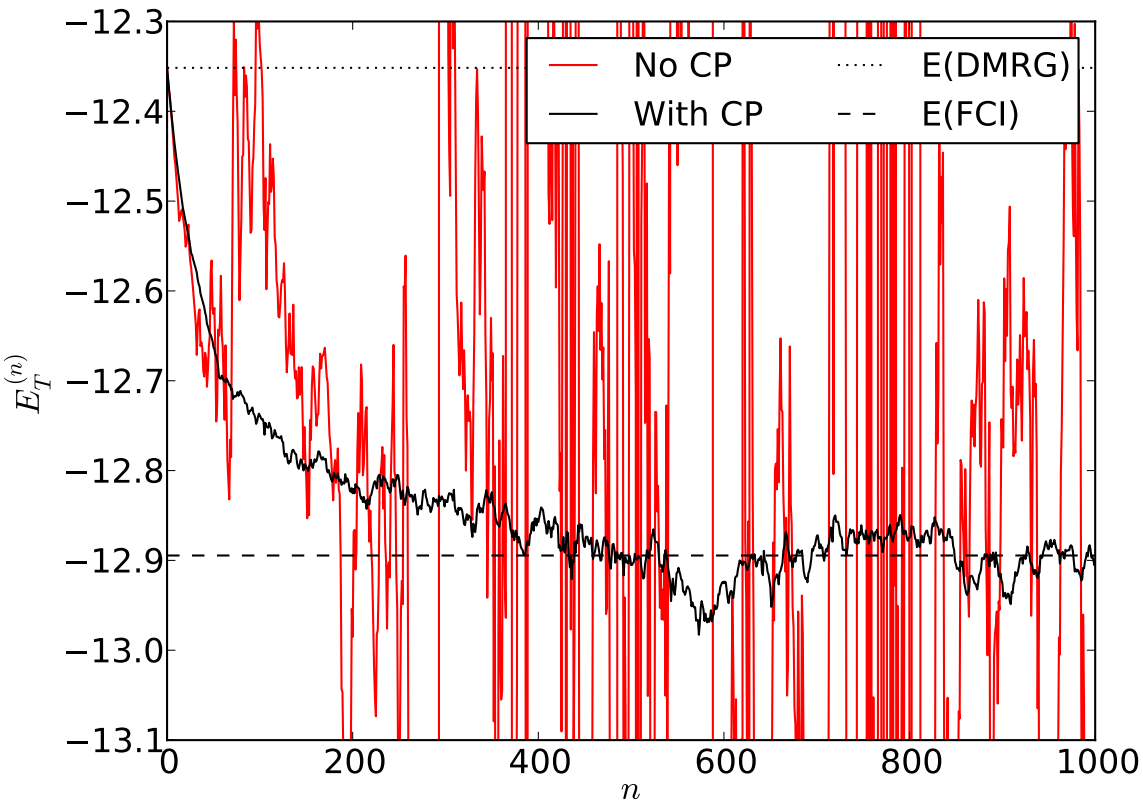
$$P(\mathbf{x}) = \frac{e^{-\mathbf{x}^2/2}}{(2\pi)^{3L/2}} \tag{21}$$

  $\hat{\mathbf{v}} = (\hat{v}_1^x, \hat{v}_1^y, \hat{v}_1^z, \hat{v}_2^x, ...)$ and L the number of lattice sites
- $\hat{B}(\mathbf{x}) \equiv \prod_i \exp\left(\sum_w \alpha_i^w \hat{S}_i^w\right)$ is a product of single-site operators
- Applying $\hat{B}(\mathbf{x})$ to an MPS does not increase its bond dimension

MPS-AFQMC: Mitigating the sign problem

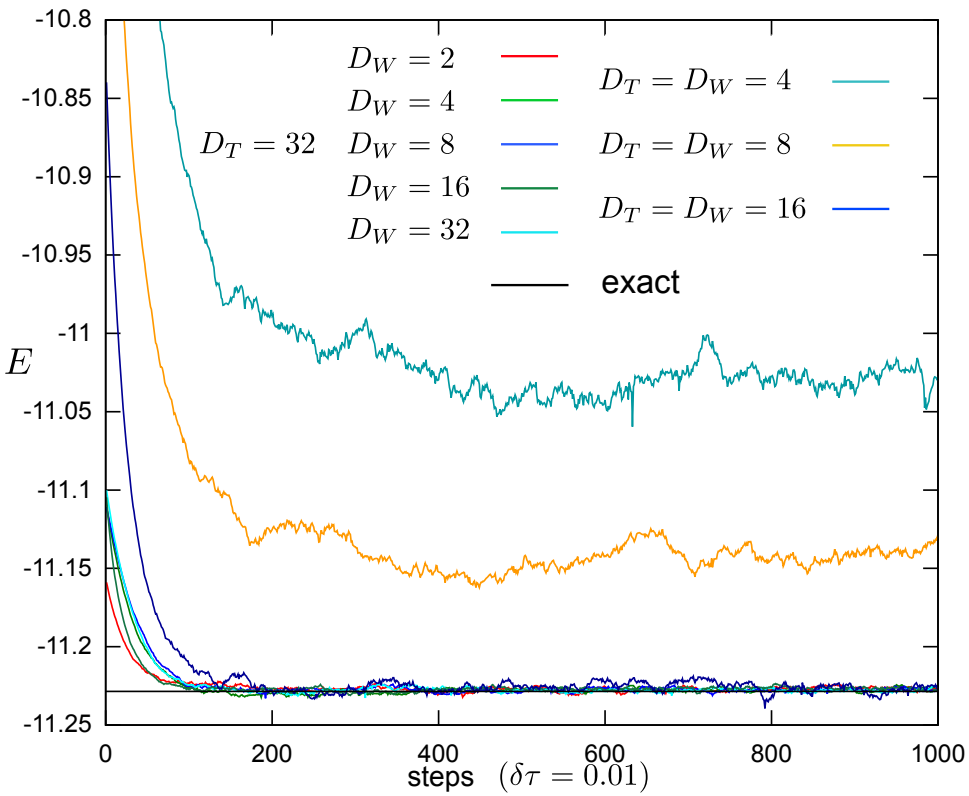
- If the operators $\hat{B}(\mathbf{x})$ allow $|\phi\rangle$ to propagate to $-|\phi\rangle$,
  then after sufficient steps $\pm|\phi\rangle$ become equally probable $\longrightarrow$ sign problem
- Variational optimization of MPS $|\Psi_T\rangle$ with bond dimension D_T: the trial wavefunction
- Constrain the walkers paths [CP] to $\langle\Psi_T \mid \phi\rangle > 0$ by importance sampling
- This introduces a systematic bias!

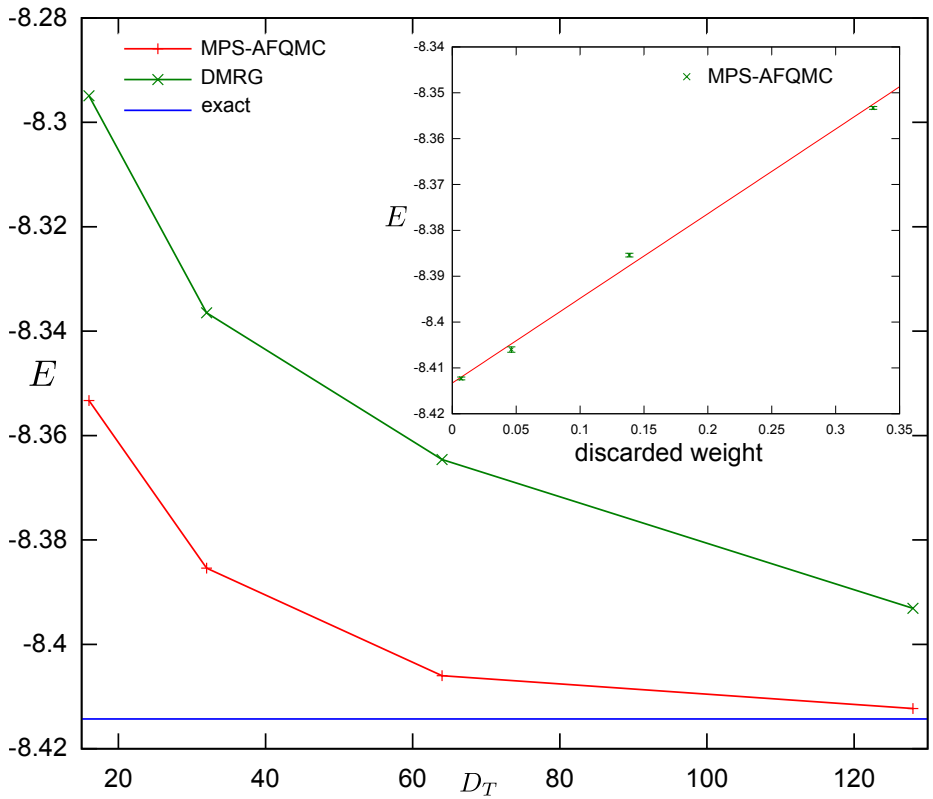
$$E_T^{(n)} = \frac{\sum_\phi \langle\Psi_T|\hat{H}|\phi\rangle}{\sum_\phi \langle\Psi_T|\phi\rangle}$$



MPS-AFQMC: Bond dimension of the walkers and the trial wavefunction

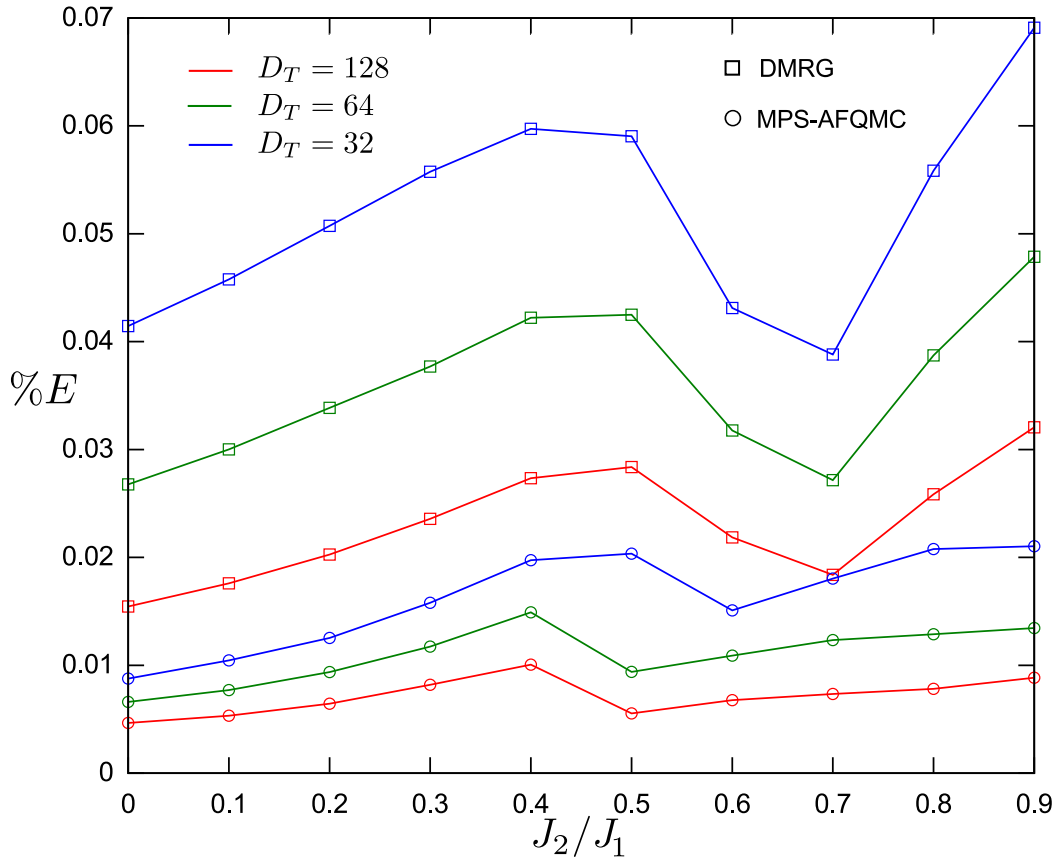
- The estimated energy does not depend on the bond dimension D_W of the walkers
- Use $D_W = 1$: Calculating $\langle \Psi_T \mid \phi \rangle$ costs $\mathcal{O}(D_W D_T^2 = D_T^2)$ instead of $\mathcal{O}(D_T^3)$
- Systematic bias due to CP can be assessed and removed by increasing D_T
- Spin-$\frac{1}{2}$ Heisenberg model on a 4×4 torus





MPS-AFQMC: Results and summary

- Spin-$\frac{1}{2}$ $J_1 J_2$-model on a 10×10 square lattice with open boundary conditions



- With Hubbard-Stratonovich walkers can have a constant bond dimension $D_W = 1$
- DMRG costs $\mathcal{O}(D_T^3)$ and MPS-AFQMC $\mathcal{O}(D_W D_T^2)$ with a large prefactor
- For higher rank tensor networks, AFQMC (PMC) can allow for a significant improvement!

Acknowledgements