COUPLED-CLUSTER METHOD TAILORED BY TENSOR-NETWORK STATES IN QUANTUM CHEMISTRY

Örs Legeza

Strongly Correlated Systems "Lendület" Research Group

Wigner Research Centre for Physics, Budapest, Hungary

in collaboration with

- Mihály Máté, Mihály Csirik (Budapest)
- Libor Veis, Jiri Pittner, Andrej Antalik, Jirka Brabec (Prague)
- Frank Neese (Germany)
- Klaas Gunst, Frank Verstraete, S. Wooters, D. van Neck (Ghent)
- Reinhold Schneider (Berlin)
- Simen Kvaal, Fabian Faulstich, Andre Laestadius (Oslo)

MOANSI 2018

Berlin, 25.10.2018

Topics to be covered

1. Tensor product factorization (mathematically exact, loop free):
   - Matrix Product State (MPS)
   - Tree Tensor Network States
   - T3NS Tensor Network States

2. Orbital entropy and Two-site mutual information

3. Basis optimization: fermionic mode transformation

4. Capturing static and dynamic correlations: DMRG-TCCSD

5. Mathematical properties of TNS-TCCSD

6. Error analysis on the N_2 molecule

Talk is based on the following references

1. *Tensor product methods and entanglement optimization for ab initio quantum chemistry*, Sz. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, Ö. Legeza, Int. J. Quant. Chem. 115:(19) 1342-1391 (2015)
2. *Coupled cluster method with single and double excitations tailored by matrix product state wave functions*, L. Veis, A. Antalik, F. Neese, Ö. Legeza, J. Pittner, J. Phys. Chem. Lett. 7, 4072 (2016)
3. *T3NS: three-legged tree tensor network states*, K. Gunst, F. Verstraete, S. Wouters, Ö. Legeza, D. Van Neck, J. Chem. Theor. Comp. **14** 2026 (2018)
4. *Coupled-Cluster Method Tailored by Tensor-Network States in Quantum Chemistry* F. M. Faulstich, A. Laestadius, S. Kvaal, Ö. Legeza, R. Schneider, arXiv:1802.05699 (2018)
5. *Numerical and Theoretical Aspects of the DMRG-TCC Method Exemplified by the Nitrogen Dimer*, F. M. Faulstich, A. Laestadius, S. Kvaal, M. Máté, M. A. Csirik, Ö. Legeza, A. Antalik, J. Brabec, L. Veis, J. Pittner, R. Schneider, preprint (2018)

DMRG provides state-of-the-art results in many fields

$$\mathcal{H} = \sum_{ij\alpha\beta} T_{ij}^{\alpha\beta} c_{i\alpha}^{\dagger} c_{j\beta} + \frac{1}{2} \sum_{ijkl\alpha\beta\gamma\delta} V_{ijkl}^{\alpha\beta\gamma\delta} c_{i\alpha}^{\dagger} c_{j\beta}^{\dagger} c_{k\gamma} c_{l\delta} \, ,$$

- T_{ij} kinetic and on-site terms, V_{ijkl} two-particle scatterings
- We consider usually lattice models in real space (DMRG)
- In quantum chemistry sites are electron orbitals (QC-DMRG)
- In UHF QC spin-dependent inetractions (UHF-QCDMRG)
- In relativistic quantum chemistry sites are spinors (4c-DMRG)
- In nuclear problems sites are proton/neutron orbitals (JDMRG)
- In k-space representation sites are momentum eigenstates (k-DMRG)
- For particles in confined potential sites $\rightarrow$ Hermite polynoms
- Major aim: to obtain the desired eigenstates of $\mathcal{H}$.

• Symmetries: Abelian and non-Abelian quantum numbers, double groups etc

• # of block states: 1 000 – 50 000. Size of Hilbert space up to 10^8.

• In ab inito DMRG the CAS size is: 50 electrons on 50 orbitals.

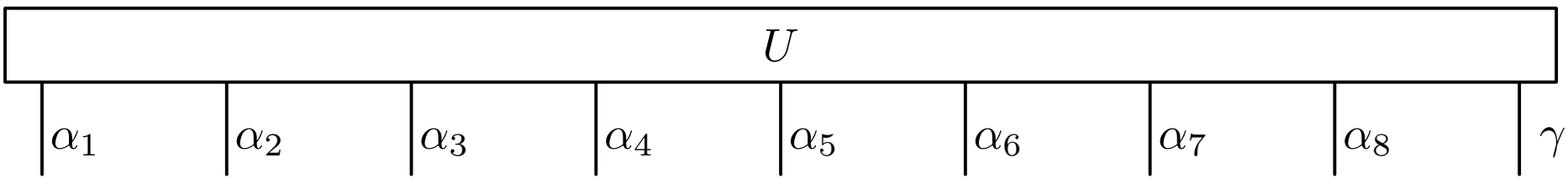
• 1-BRDM and 2-BRDM can be extracted.

Tensor product approximation

State vector of a quantum system in the discrete tensor product spaces

$$|\Psi_\gamma\rangle = \sum_{\alpha_1=1}^{n_1} \dots \sum_{\alpha_d=1}^{n_d} U(\alpha_1, \dots, \alpha_d, \gamma)\, |\alpha_1\rangle \otimes \cdots \otimes |\alpha_d\rangle \in \bigotimes_{i=1}^{d} \Lambda_i := \bigotimes_{i=1}^{d} \mathbf{C}^{n_i} \ ,$$

where $span\{|\alpha_i\rangle : \alpha_i = 1, \dots, n_i\} = \Lambda_i = \mathbf{C}^{n_i}$ and $\gamma = 1, \dots, m$.



In a spin-1/2 model $\alpha_i \in \{\downarrow, \uparrow\}$.

In a spin-1/2 fermionic model $\alpha_i \in \{0, \downarrow, \uparrow, \uparrow\downarrow\}$.

$\dim \mathcal{H}_d = \mathcal{O}(n^d)$ Curse of dimensionality!

Tucker representation or approximation

One is looking good or even optimal bases

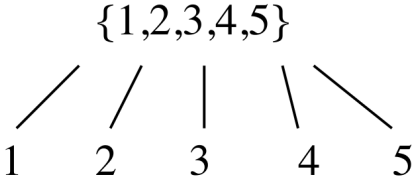
$$\{|\xi^i_{m_i}\rangle : m_i = 1, \ldots, r_i\} \simeq \{\alpha_i \mapsto \xi_i(m_i, \alpha_i) : m_i = 1, \ldots, r_i\}$$

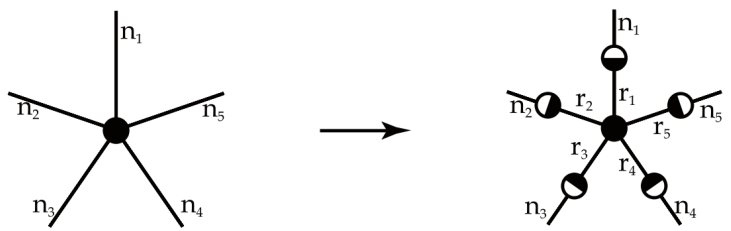
of size $r_i \leq n_i$, in each coordinate direction α_i, $i = 1, \ldots, d$, give the representation (or approximation)

$$|\Psi_y\rangle = \sum_{m_1=1}^{r_1} \cdots \sum_{m_d=1}^{r_d} C(m_1, \ldots, m_d, y) |\xi^1_{m_1}\rangle \otimes \cdots \otimes |\xi^d_{m_d}\rangle \ , \ y = 1, \ldots, m \ .$$

or in terms of coefficients

$$U(\alpha_1, \ldots, \alpha_d, y) = \sum_{m_1=1}^{r_1} \cdots \sum_{m_d=1}^{r_d} C(m_1, \ldots, m_d, y) \xi_1(\alpha_1, m_1) \ldots \xi_d(\alpha_d, m_d)$$





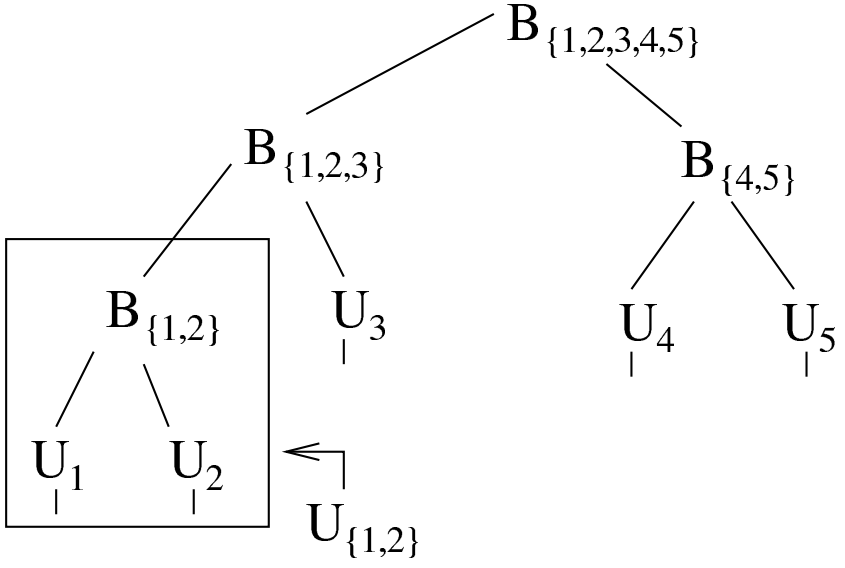
big reduction from $\mathcal{O}(mn^d)$ to $\mathcal{O}(rnd + mr^d)$, but still scales exponentially with d.

Hierarchical tensor (HT) approximation

We pursue not performing this idea in one step, but proceed in a hierarchical way

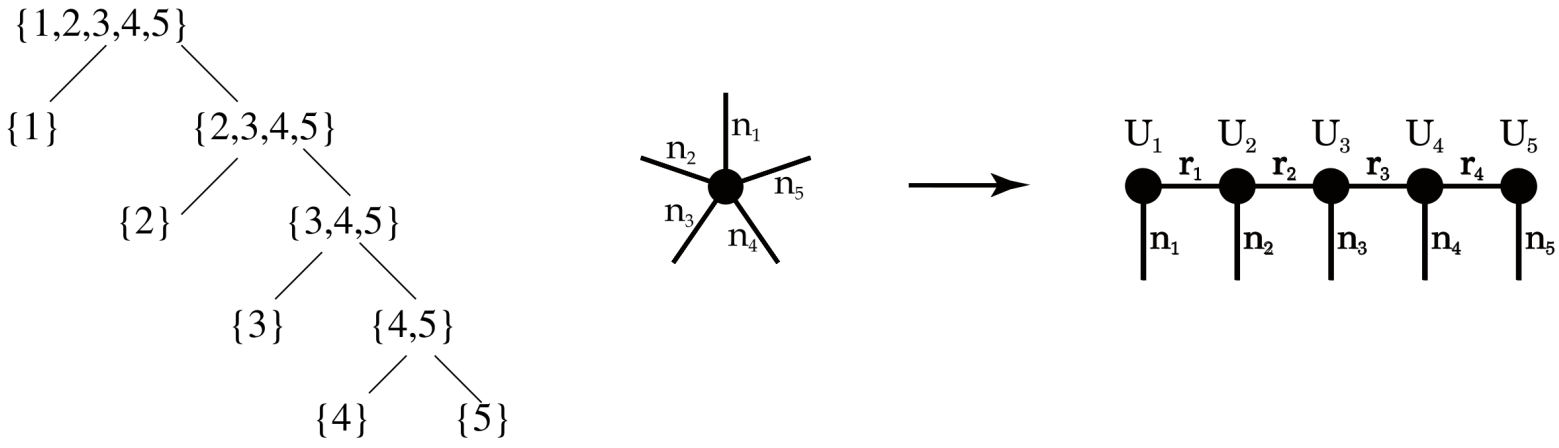
For the approximation of U, we may need in $V_1 \otimes V_2$ only a subspace $V_{\{1,2\}} \subset V_1 \otimes V_2$ with dimension $r_1 < n_1 n_2$, this is defined through a new basis given in the Tucker representation as

$$|\xi^{\{1,2\}}_{m_{\{1,2\}}}\rangle = \sum_{m_1=1}^{n_1} \sum_{m_2=1}^{n_2} U_{\{1,2\}}(m_{\{1,2\}}, \alpha_1, \alpha_2)\, |\alpha_1\rangle \otimes |\alpha_2\rangle \ .$$



Various possibilities to build partition trees: e.g. NRG

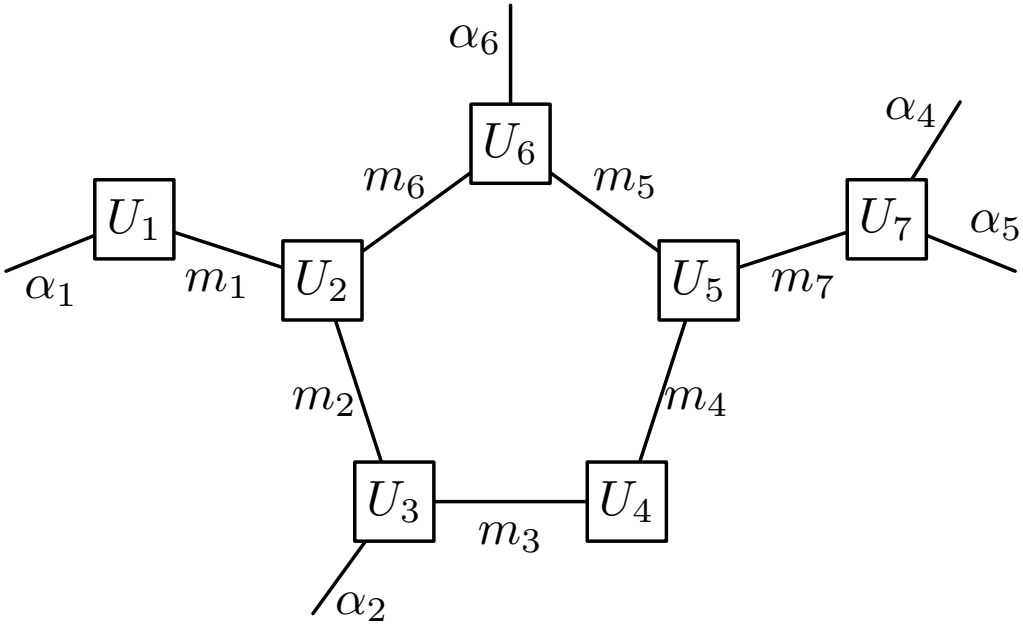
Numerical Renormalization group method, (Wilson, 1975)



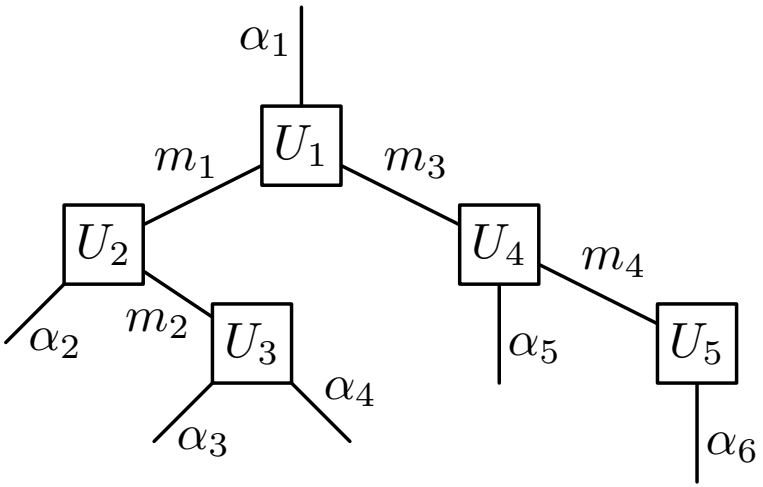
For a concrete problem, one has to choose an appropriate tree. This choice has a tremendous influence onto the efficiency of the hierarchical tensor representation.

The optimal ranks of the tensors, $r_i \ll n^d$, are determined by the Schmidt-decomposition$\rightarrow$ strong connection to quantum information theory

Tensor product representation



A general tensor network representation of a tensor of order 5.



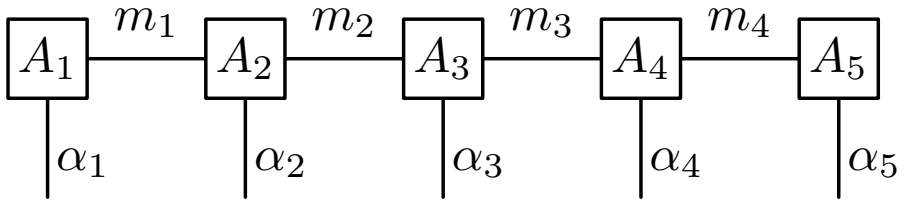
An arbitrary example of a tensor tree (loop free).

Matrix product state (MPS) representation

The tensor U is given element-wise as

$$U(\alpha_1, \ldots, \alpha_d) = \sum_{m_1=1}^{r_1} \ldots \sum_{m_{d-1}=1}^{r_{d-1}} A_1(\alpha_1, m_1) A_2(m_1, \alpha_2, m_2) \cdots A_d(m_{d-1}, \alpha_d). \tag{1}$$

We get d component tensors of order 2 or 3.



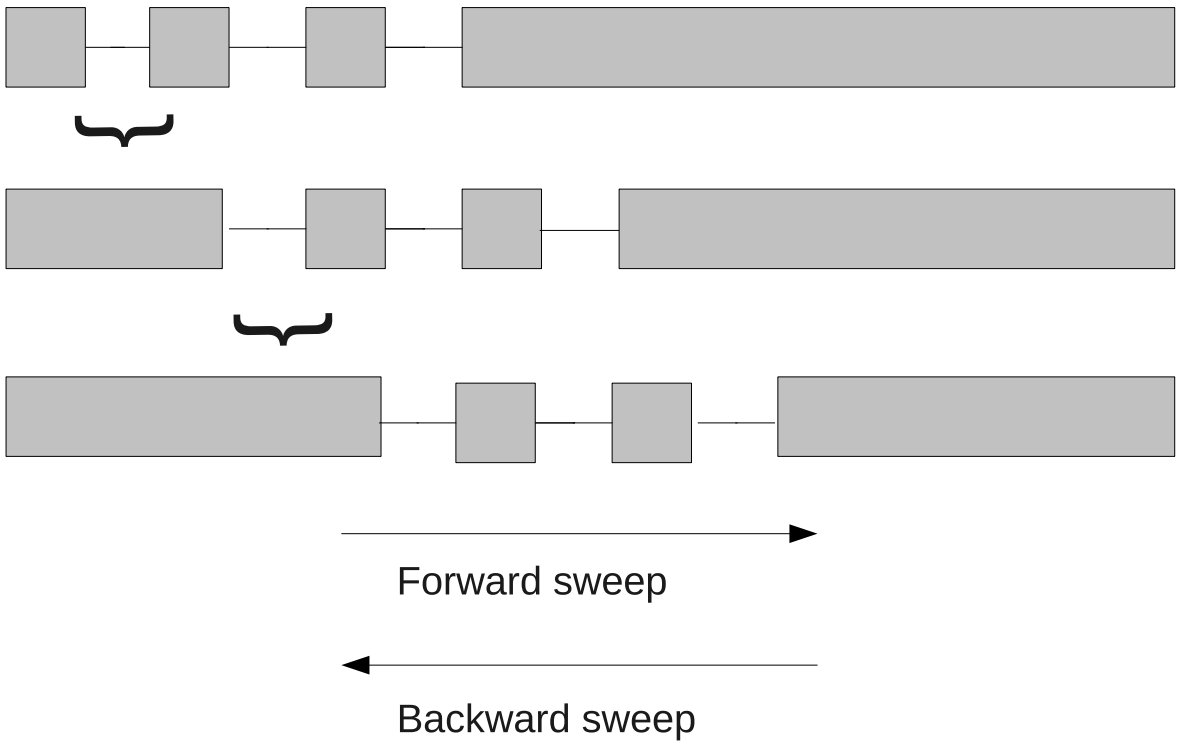
A tensor of order 5 in Matrix Product State (MPS) representation.
Also know as Tensor Train (TT).
This yields a chain of matrix products:

$$U(\alpha_1, \ldots, \alpha_d) = \mathbf{A}_1(\alpha_1)\mathbf{A}_2(\alpha_2) \cdots \mathbf{A}_{d-1}(\alpha_{d-1})\mathbf{A}_d(\alpha_d) \tag{2}$$

with $[\mathbf{A}_i(\alpha_i)]_{m_{i-1}, m_i} := A_i(m_{i-1}, \alpha_i, m_i) \in \mathbb{C}^{r_{i-1} \times r_i}$.

Redundancy:
$U(\alpha_1, \ldots, \alpha_d) = \mathbf{A}_1(\alpha_1)\mathbf{G}\mathbf{G}^{-1}\mathbf{A}_2(\alpha_2) \cdots \mathbf{A}_{d-1}(\alpha_{d-1})\mathbf{A}_d(\alpha_d)$

Density matrix renormalization group wavefunction White (1992)



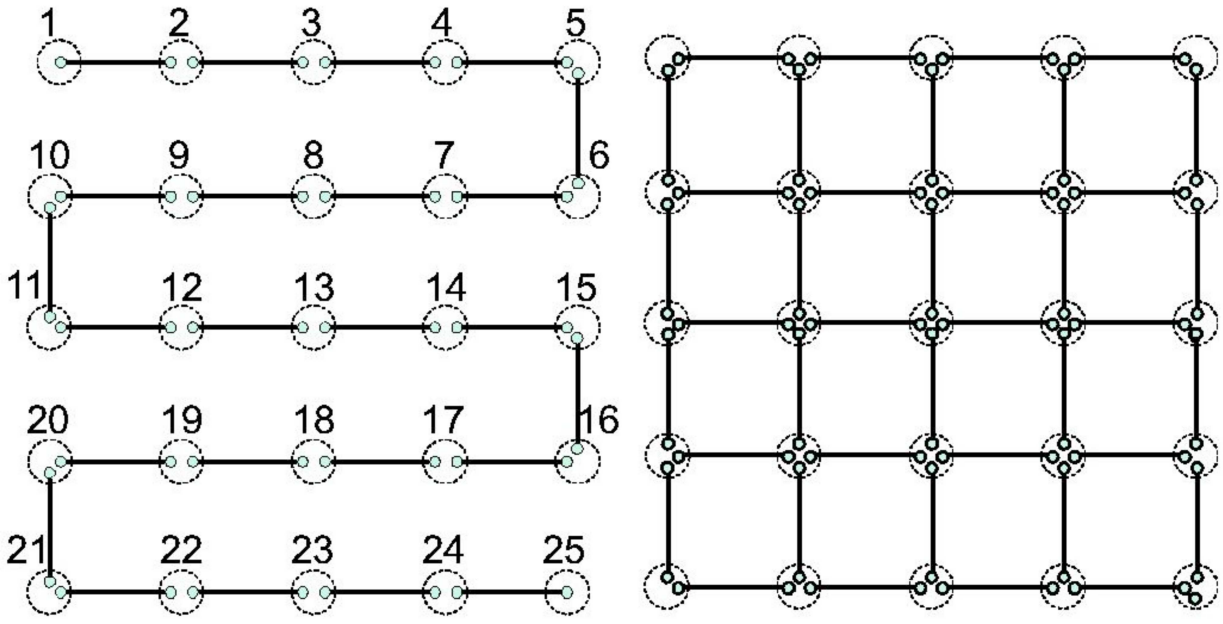
$$|\Psi_{\mathrm{TG}}\rangle = \sum_{\alpha_l \alpha_{l+1} \alpha_{l+2} \alpha_r} \psi_{\alpha_l \alpha_{l+1} \alpha_{l+2} \alpha_r} |\phi^{(l)}_{\alpha_l}\rangle \otimes |\phi^{(s_l)}_{\alpha_{l+1}}\rangle \otimes |\phi^{(s_r)}_{\alpha_{l+2}}\rangle \otimes |\phi^{(r)}_{\alpha_r}\rangle$$

where $\psi_{\alpha_l \alpha_{l+1} \alpha_{l+2} \alpha_r}$ coefficients (4-index tensor) are determined by an iterative diagonalization of the superblock Hamiltonian.

DMRG algorithm provides the optimized set of A_i matrices through a series of unitary transformation based on the singular value decomposition (SVD) theorem by sweeping through the network.

Extension of MPS to higher dimensional cases: PEPS

- For 2D systems MPS representation is not optimal
- Short range interactions become also long range

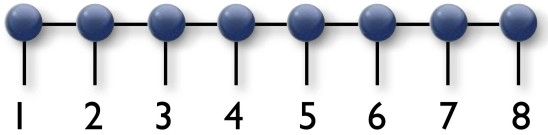


- Entanglement in all 4 direction $\rightarrow$ tensor product states needed!
- Use tensors $A^i[\alpha]_{m_1,m_2,m_3,m_4}$
- Projected Entangled-Pair State (PEPS)

Various tensor methods exist:

1D MPS

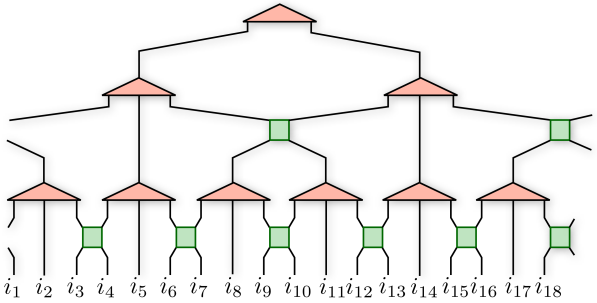
Matrix-product state



White, Östlund, Rommer

1D MERA

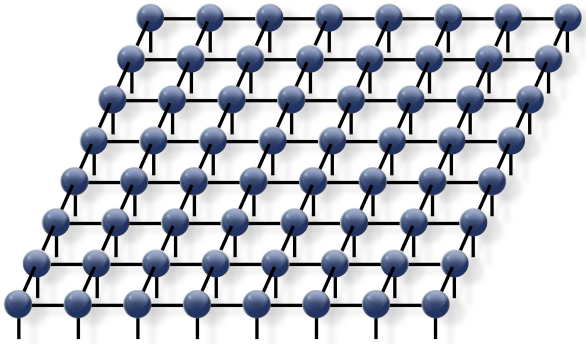
Multi-scale entanglement renormalization ansatz



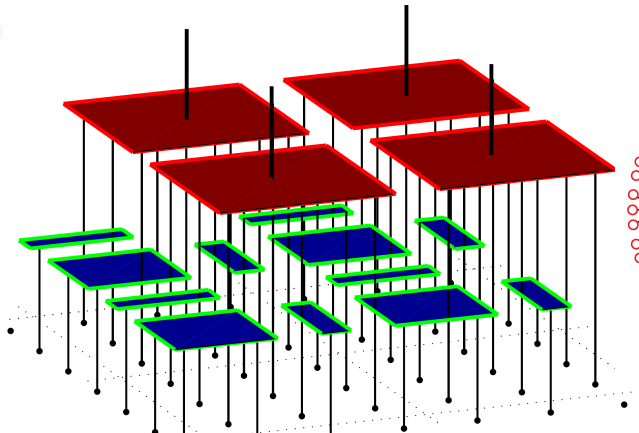
1D TTNS

Tree tensor network state Vidal, Corboz

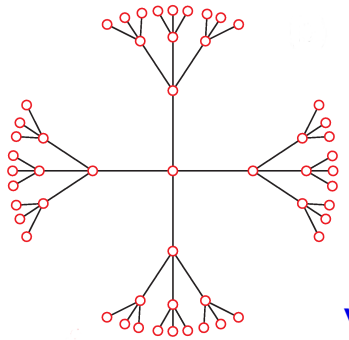
2D PEPS



2D Mera



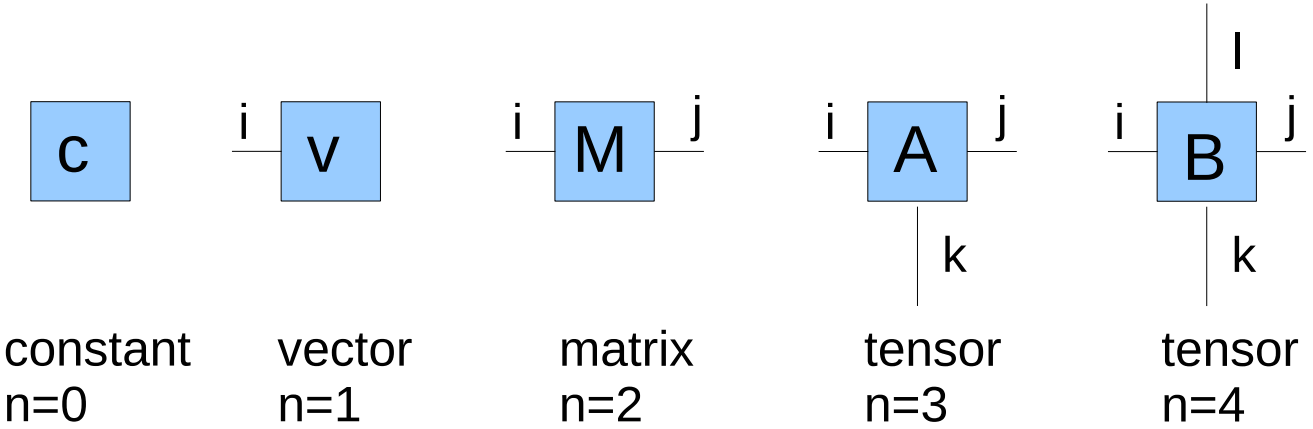
2D Tree-TNS



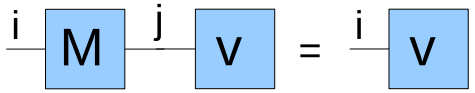
Vidal, Corboz

A little tensor algebra:

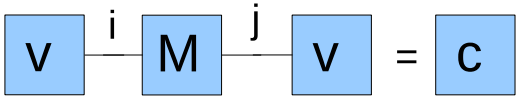
- Building blocks of the networks: tensors with n open legs



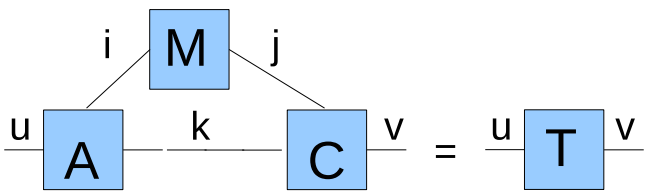
- In the networks connected lines correspond to contraction: sum over related indicies:
- $\sum_{ij} M_{ij} v_j = u_i$



- $\sum_{ij} u_i M_{ij} v_j = c$



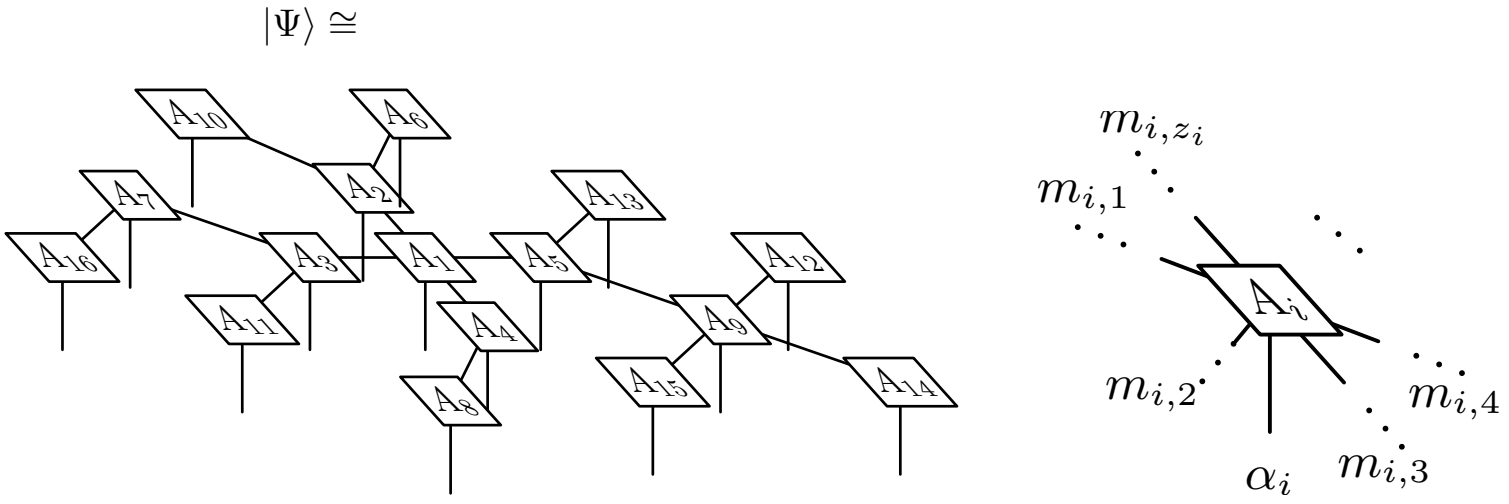
- $\sum_{ijk} A_{uik} M_{ij} C_{vjk} = T_{uv}$



- Order of contraction is important

Higher dimensional networks (Ex.: Tree-TNS)

Corboz, Vidal (2009), Murg, Verstraete, Ö.L, Noack (2010, 2014), Nakatani, Chan (2013)

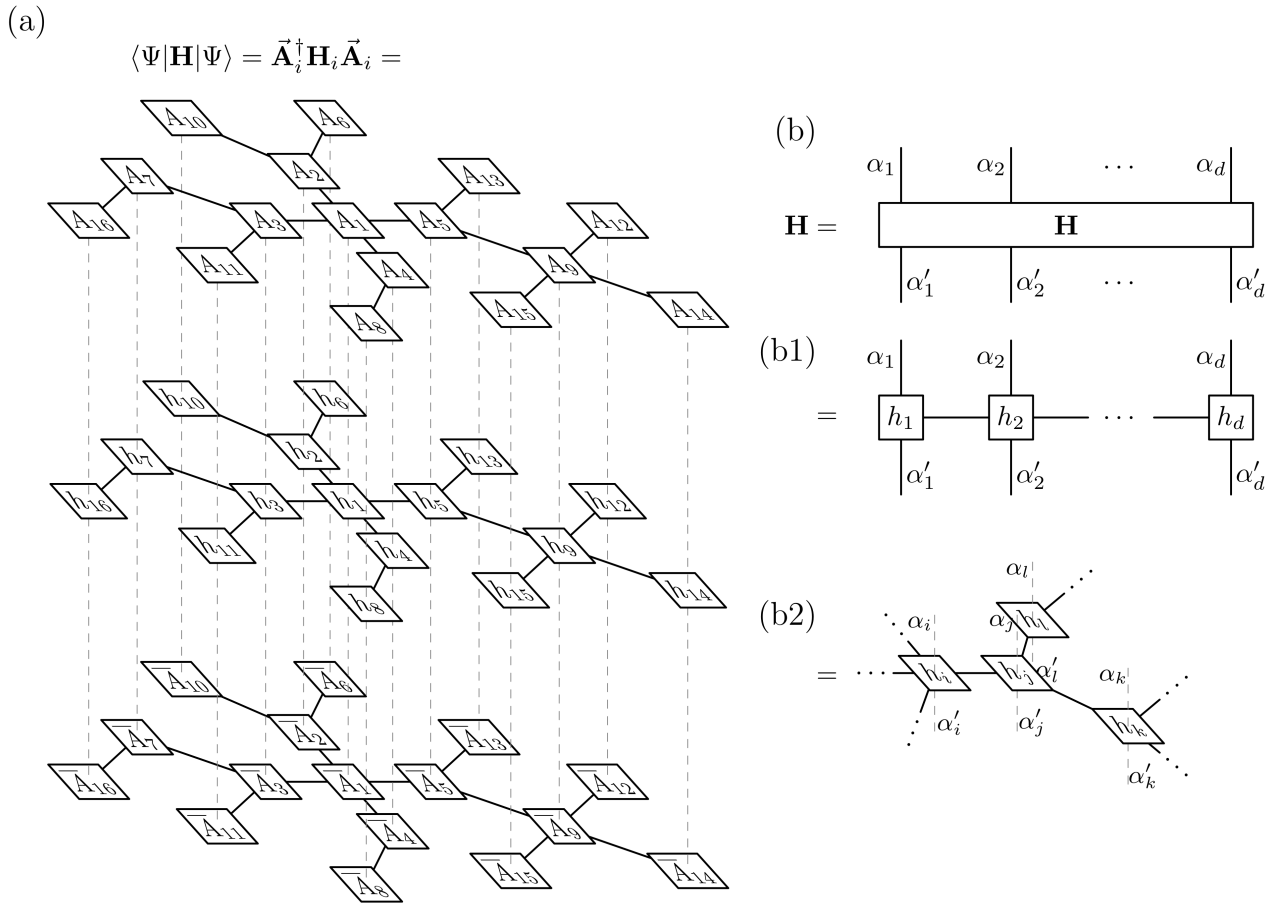


Schematic plot of a higher dimensional network, for example, the tree tensor network state (TTNS). Each node is represented by a tensor A_i of order $z_i + 1$, with z_i is a site dependent coordination number.

The network supposed to reflect the entanglement structure of the system as much as possible.

Maximal distance between two sites, 2Δ, scales logarithmically with d for $z > 2$.

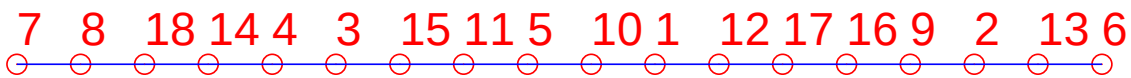
Decomposition of the Hamiltonian as TTNO



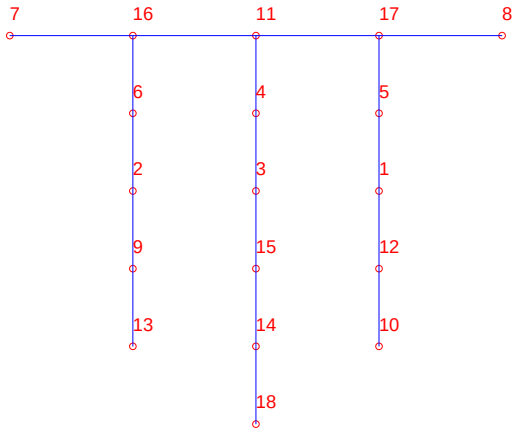
(a) expectation value $\langle\Psi|\mathbf{H}|\Psi\rangle$ with respect to the TTNS
(b) The Hamiltonian $\mathbf{H}$, represented as TTNO of component tensors $\mathbf{h}_i$ in the middle.
(b1) decomposition of the Hamiltonian as MPO
(b2) decomposition of the Hamiltonian as TTNO

Variable tensor orders and convergence properties

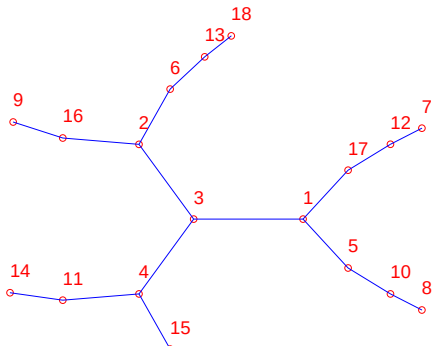
(a) Network-1

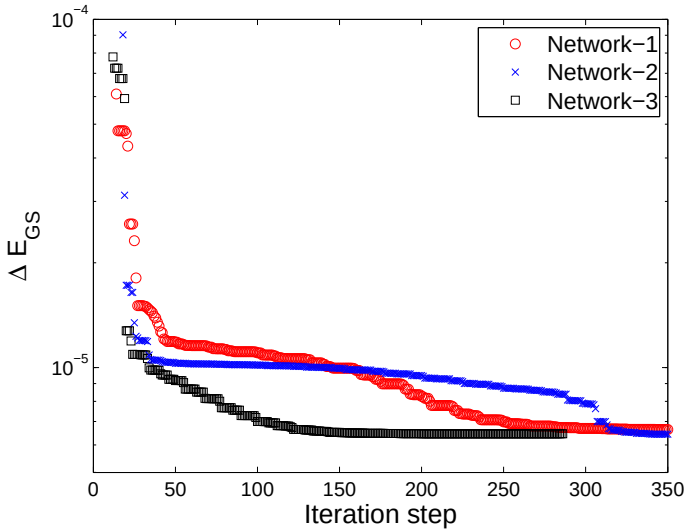


(b) Network-2



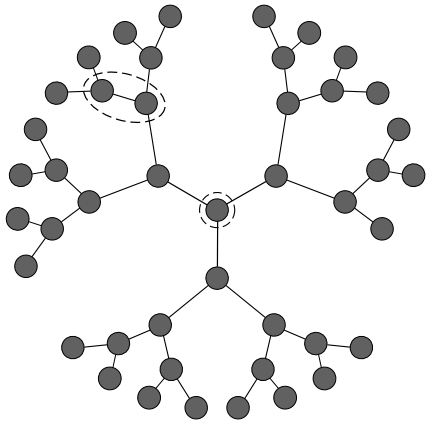
(c) Network-3



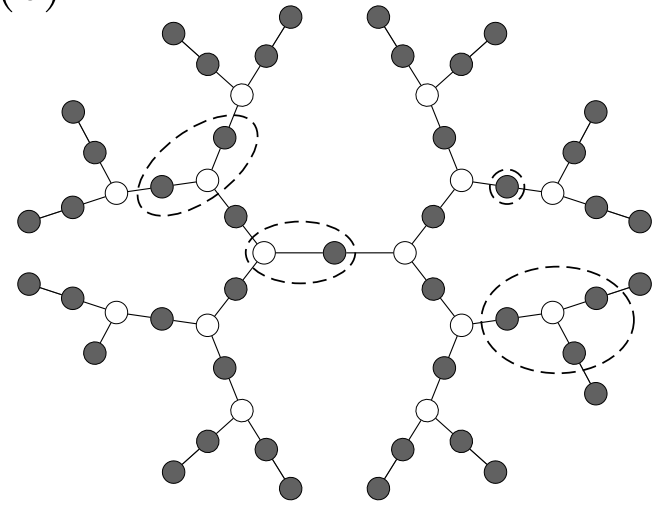


T3NS a new tensor format Gunst, Verstraete, Wooters, Ö.L., van Neck (2018)

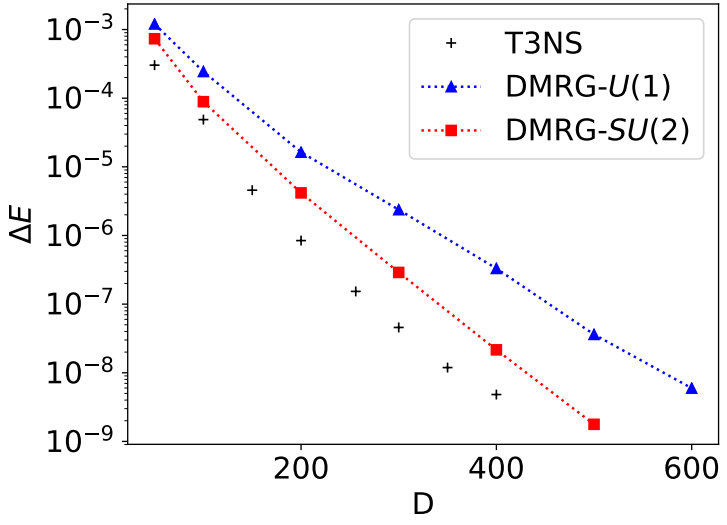
(a)



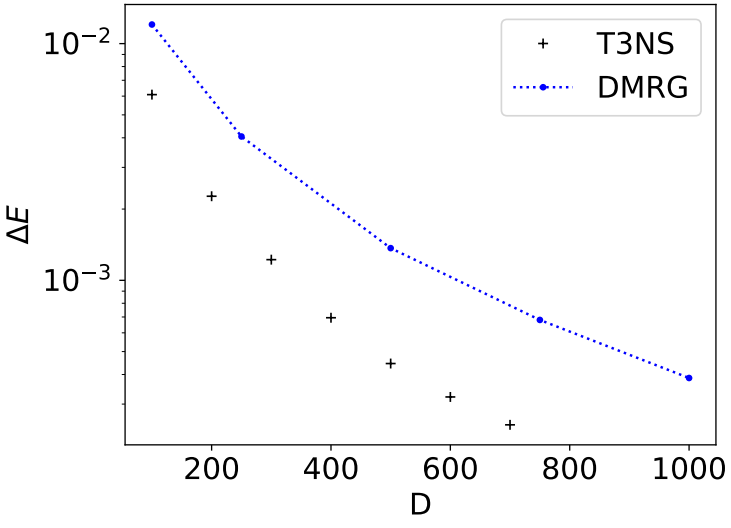
(b)



LiF



N_2



Example: $[Cu_2O_2]^{2+}$

Ref. Method	$E_{\mathrm{bisoxo}}[\mathrm{E_h}]$	$E_{\mathrm{peroxo}}[\mathrm{E_h}]$	ΔE [kcal/mol]
37 CASSCF(16,14)			0.2
37 CASPT2(16,14)			1.4
38 RASPT2(24,28)			28.7
Some previously published DMRG energies			
40 DMRG(32,62)[2400]			35.6
41 DMRG(28,32)[2048]-SCF/CT			27.0
43 DMRG(32,28)[4000]			21.8
44 DMRG(24,24)[1500]-SCF*			35.1
44 DMRG(24,24)[1500]-CASPT2*			23.2
39 DMRG(26,44)[800]	-541.46779	-541.49731	18.5
42 DMRG(26,44)[128]	-541.47308	-541.51470	26.1
33 DMRG(26,44)[256/1024/10^{-5}]†	-541.53853	-541.58114	26.7
T3NS calculations			
T3NS(26,44)[50]	-541.48773	-541.56999	51.6
T3NS(26,44)[100]	-541.52352	-541.57166	30.2
T3NS(26,44)[200]	-541.53284	-541.57717	27.8
T3NS(26,44)[300]	-541.53556	-541.57966	27.7
T3NS(26,44)[500]	-541.53820	-541.58094	26.8

Resource requirements and complexity

	DMRG	T3NS
CPU time:	$\mathcal{O}\left(k^4D^2 + \underline{k^3D^3}\right)$	$\mathcal{O}\left(k^5D^2 + \underline{k^3D^4}\right)$
Memory:	$\mathcal{O}\left(k^2D^2\right)$	$\mathcal{O}\left(k^2D^2 + kD^3\right)$
Disk:	$\mathcal{O}\left(k^3D^2\right)$	$\mathcal{O}\left(k^3D^2 + kD^3\right)$

- k : number of orbitals
- D: bond dimension
- The underlined terms correspond with the complexity of the most intensive part of the algorithm, i.e. the matrix-vector product used in the iterative solver.

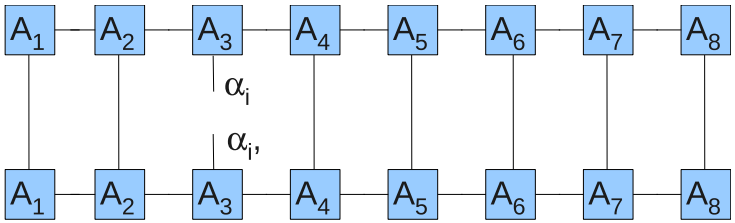
One- (ρ_i) and two-orbital ($\rho_{i,j}$) reduced density matrix

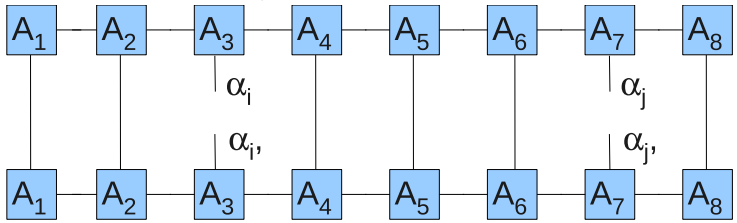
$$|\psi\rangle = \sum_{\alpha_1,\ldots,\alpha_N} C_{\alpha_1,\ldots,\alpha_N} |\alpha_1 \ldots \alpha_N\rangle \ ,$$

- $\rho_{i,j}$ is calculated by taking the trace of $|\Psi\rangle\langle\Psi|$ over all local bases except for α_i and α_j, the bases of sites i and j, i.e.,

$$\rho_{i,j}([\alpha_i, \alpha_j], [\alpha_i', \alpha_j']) = \sum_{\substack{\alpha_1,\ldots,\cancel{\alpha_i},\ldots,\\ \cancel{\alpha_j},\ldots,\alpha_N}} C_{\alpha_1,\ldots,\alpha_i,\ldots,\alpha_j,\ldots,\alpha_N} \ C^*_{\alpha_1,\ldots,\alpha_i',\ldots,\alpha_j',\ldots,,\alpha_N} \ .$$

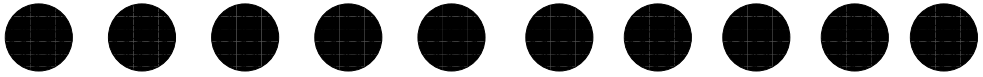
- In the MPS representation, calculation of ρ_{ij} corresponds to the contraction of the network except at sites i and j.



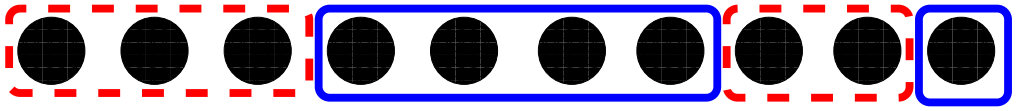


- This can be decomposed as a sum of projector operators based on the free variables α_i and α_j.
- ρ_i and $\rho_{i,j}$ can be constructed from operators describing transitions between single-site basis states.

Mutual information: classical and quantum correlations

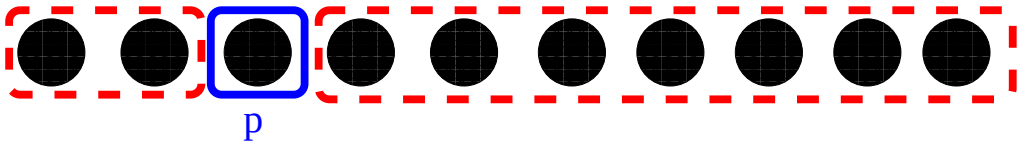


$$\varrho = |\psi\rangle\langle\psi|$$



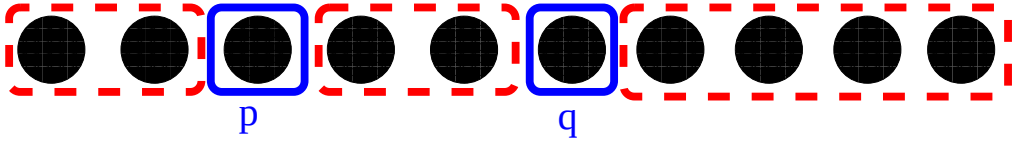
$$\varrho^B = \mathrm{Tr}_\mathrm{A}\varrho$$

$$S^B = -\mathrm{Tr}(\varrho^\mathrm{B}\ln\varrho^\mathrm{B})$$



$$\varrho^p \Rightarrow S^p$$

S^p describes the entanglement of site p with the rest of the system.



$$\varrho^{p,q} \Rightarrow S^{p,q}$$

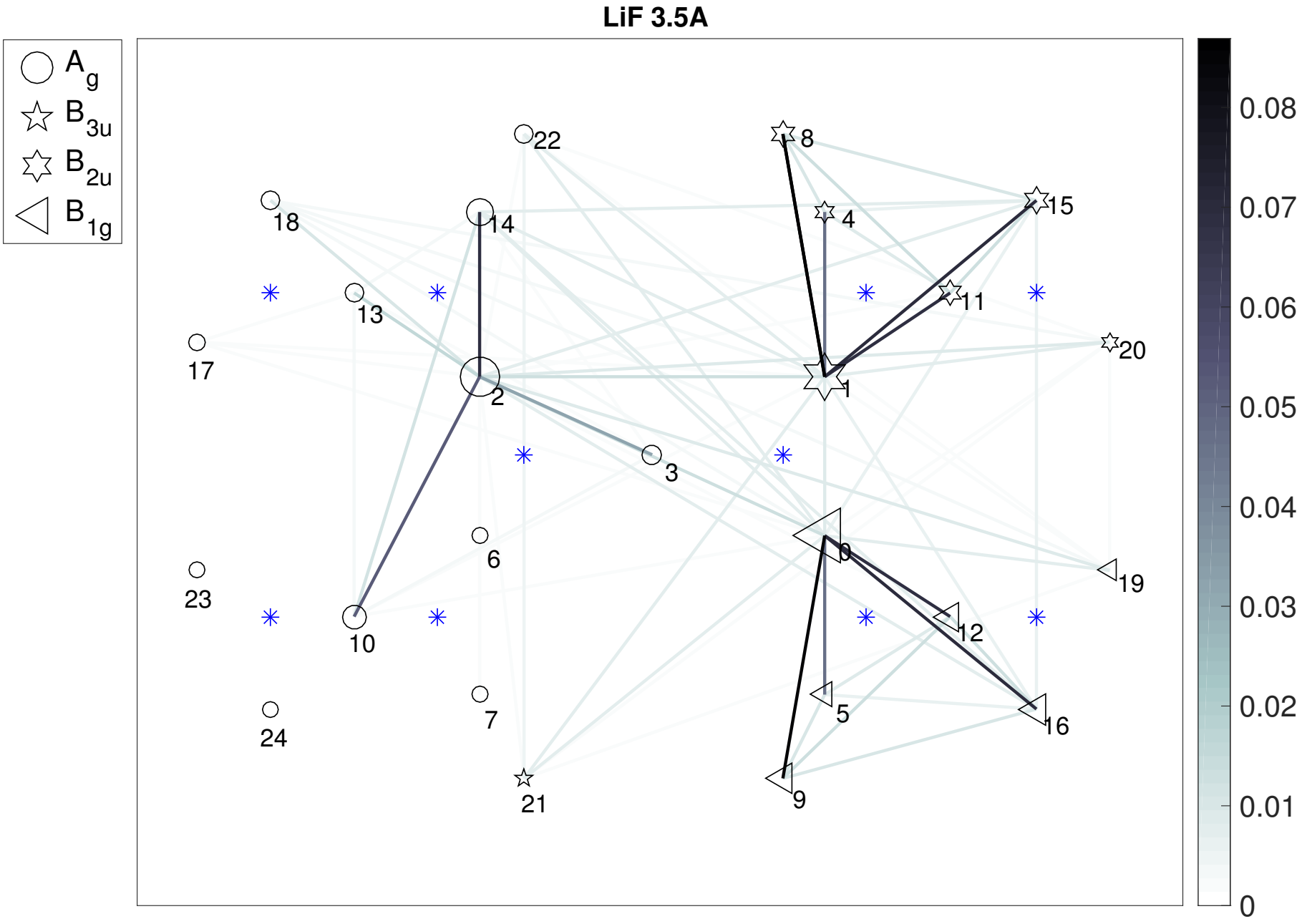
$S^{p,q}$ describes the entanglement of orbital p and q with the rest of the system.
$I^{p,q}$ describes the mutual information between orbital p and q

$$I^{p,q} = S^p + S^q - S^{p,q}$$

Ö.L., Sólyom, PRB (2003): Quantum Chemistry,
Ö.L., Sólyom, PRL (2005): quantum phase transitions (QPT) with $q = p + 1$.

Rissler, White, Noack, ECP (2005): Quantum chemistry, arbitrary p and q.

Network optimization by the mutual information

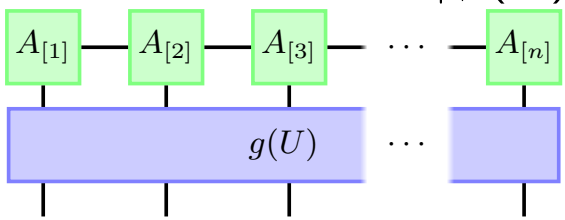


Redefinition of the fermionic modes by a linear transformation

- Linear transformations of a set of fermionic annihilation operators $\{c_i\}$ to a new set $\{d_i\}$ satisfying the canonical anti-commutation relations:

$$c_i = \sum_{j=1}^{Np} U_{i,j} d_j, \quad p \text{ denotes the number of different fermion species}$$

- Under this change of basis a state vector $|\psi(U)\rangle = G(U)|\psi(\mathbb{1})\rangle$



- Denoting the Hamiltonian written in terms of the transformed modes by $H(U) = G(U)^\dagger H G(U)$, we are interested in the solutions of

$$(U_{\text{opt}}, |\psi_{\text{opt}}\rangle) = \underset{\substack{U \in U(Np), \\ |\psi\rangle \in \mathcal{M}_{D_{\max}}}}{\text{argmin}} \langle\psi|H(U)|\psi\rangle.$$

- The global basis change is composed of local unitaries solutions of

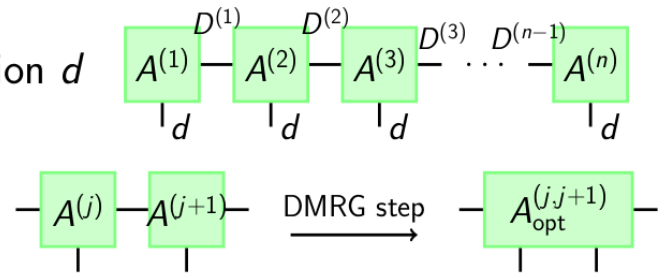
$$U_{\text{opt}}^{\text{loc}} = \text{argmin}_{U \in V} f_j\big(|\psi(\mathbb{1}_j \oplus U \oplus \mathbb{1}_{N-j-2})\rangle\big),$$

cost function $f_j^{(1)}(|\psi\rangle) = ||\Sigma_\psi^j||_1$ where Σ_ψ^j denotes the Schmidt spectrum of $|\psi\rangle$ for a bipartiting cut between sites j and $j+1$.

Local mode transformation: black-box tool to improve basis

Krumnow, Veis, Ö. L., Eisert, 2014-2016

- Perform updates **iteratively and adaptively**, both in the MPS ansatz and in mode transformations.
- Consider a matrix-product state with physical dimension d and maximal bond dimension $D_{\text{max}} = \max\{D^{(j)}\}$.
- For given $j \in \{1, \ldots, n-1\}$, minimize the energy by jointly optimizing the tensors $A^{(j)} \in \mathbb{C}^{D^{(j-1)} \times D^{(j)} \times d}$ and $A^{(j+1)} \in \mathbb{C}^{D^{(j)} \times D^{(j+1)} \times d}$ at sites j and $j+1$

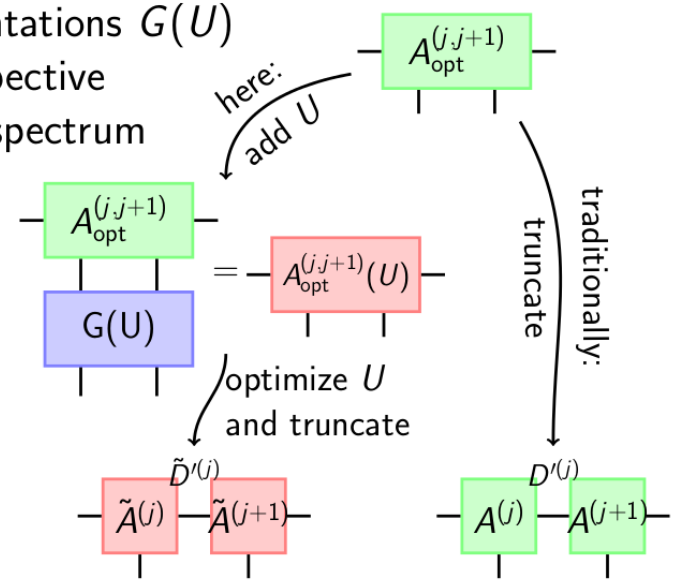


- Jointly update $A^{(j)}, A^{(j+1)}$ with Hilbert space representations $G(U)$ of mode transformations $U \in U(2 \log_2 d)$ on the respective physical legs of the tensors, optimizing the Schmidt-spectrum of $A^{(j,j+1)}_{\text{opt}}(U)$ over the cut $j, j+1$ and truncate.
- Update the operators with $U_{\text{global}} := \mathbb{1} \oplus U \oplus \mathbb{1}$ e.g. the **Hamiltonian**

$$H \mapsto \tilde{H} := G(U_{\text{global}})\, H\, G^\dagger(U_{\text{global}})$$

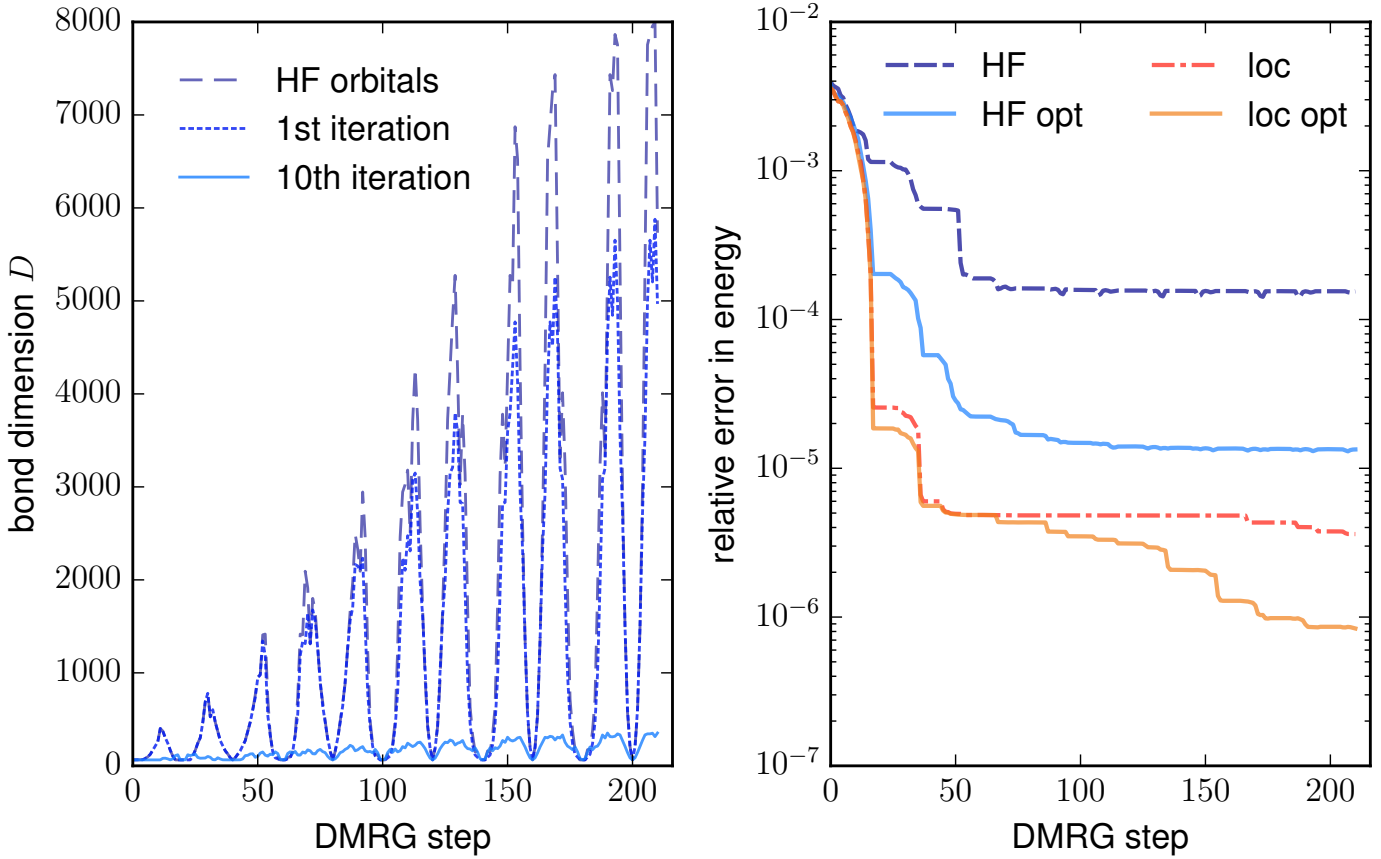
exploiting their second quantized representation

$$H(T, V) \mapsto \tilde{H} = H(\tilde{T}, \tilde{V})$$
$$T \mapsto \tilde{T} := U_{\text{global}}\, T\, U^\dagger_{\text{global}}$$
$$V \mapsto \tilde{V} := (U_{\text{global}} \otimes U_{\text{global}})\, V\, (U^\dagger_{\text{global}} \otimes U^\dagger_{\text{global}})$$



- Go to next site $j \mapsto j \pm 1$ and iterate.
- Build up a global non-trivial mode transformation by consecutive local mode transformations with overlapping support
- At some point, fix the basis (which has now been optimised to the MPS ansatz and not Renyi entropic qualifiers) and perform state-of-the art DMRG with large bond dimension.

Large-scale DMRG results (Ex.: Be_6 ring)

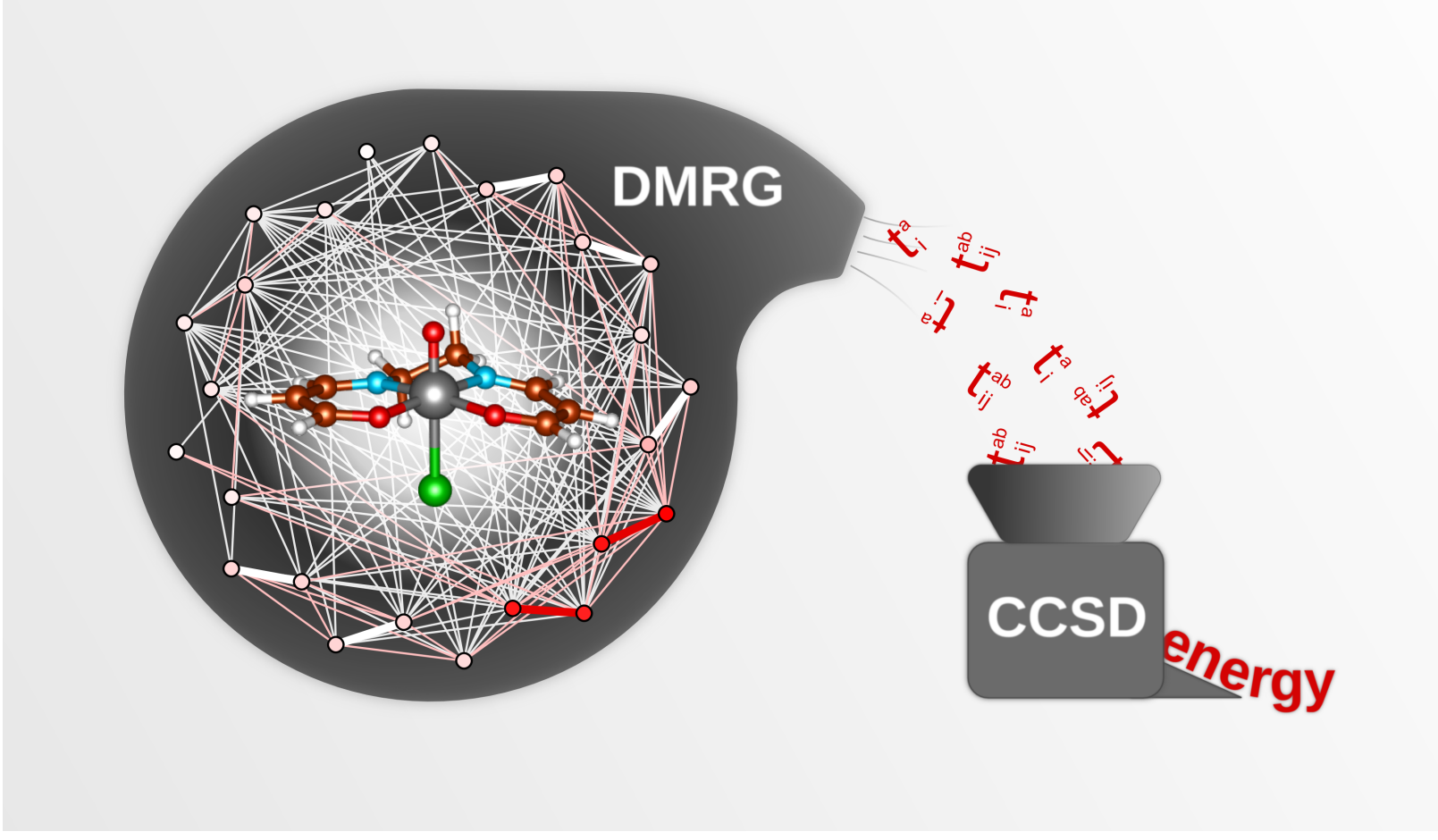


Left panel: bond dimension needed for a bounded truncation error $\epsilon_{trc} \leq 10^{-6}$ and $D_{\text{min}} = 64$ when starting in the HF basis.
Right panel: the relative error in energy $(\langle\psi|H|\psi\rangle - E_0)/E_0$ obtained by calculations with $D_{\text{max}} = 256$.
E_0 was obtained from a calculation with $D_{\text{max}} = 2048$ in the localized basis.

Coupled cluster method with single and double excitations tailored by matrix product state wave functions

L. Veis, A. Antalik, F. Neese, Ö.L., J. Pittner (2016)

- Efficient treatment of static and dynamic correlations based on TCCSD method of Bartlett [Kinoshita, Hino, and Bartlett, JCP 123, 074106(2005)]



Tailored coupled clusters

- Formally single reference theory, Fermi vacuum is a single determinant
- Split-amplitude ansatz

$$|\Psi_{\mathrm{TCC}}\rangle = e^{\mathcal{T}} |\Psi_{\mathrm{ref}}\rangle = e^{\mathcal{T}^{\mathrm{ext}} + \mathcal{T}^{\mathrm{CAS}}} |\Psi_{\mathrm{ref}}\rangle$$

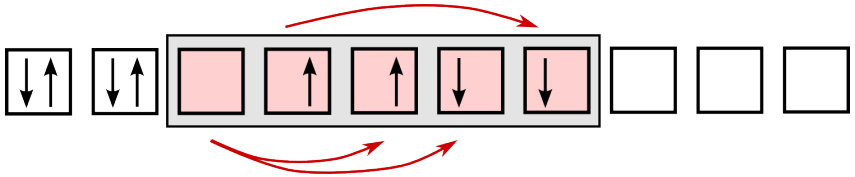
- $\mathcal{T}^{\mathrm{CAS}}$
  - amplitudes extracted from DMRG (CASCI) calculation
  - frozen during CC calculation
  - account for static correlation
- $\mathcal{T}^{\mathrm{ext}}$
  - determined through the usual CC
  - account for dynamic correlation

$$\begin{aligned} |\Psi_{\mathrm{TCCSD}}\rangle &= e^{\left(\mathcal{T}_1^{\mathrm{ext}} + \mathcal{T}_2^{\mathrm{ext}}\right)} e^{\left(\mathcal{T}_1^{\mathrm{CAS}} + \mathcal{T}_2^{\mathrm{CAS}}\right)} |\Psi_{\mathrm{ref}}\rangle \\ &\approx e^{\left(\mathcal{T}_1^{\mathrm{ext}} + \mathcal{T}_2^{\mathrm{ext}}\right)} |\Psi_{\mathrm{CASCI}}\rangle \end{aligned}$$

- Requires only small modifications of the CC code

CCSD tailored by MPS wave functions

1. Small active space DMRG calculation



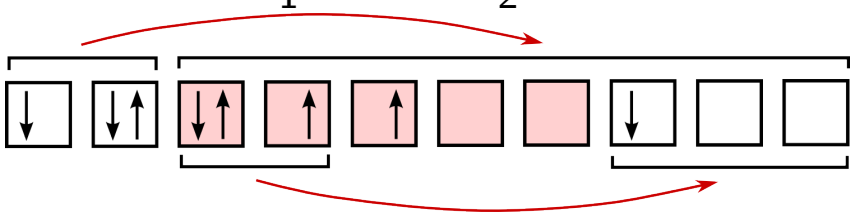
2. Acquisition of CI coefficients by efficient contraction of MPS w.f. (in two-site form)

$$|\Psi_{\mathrm{MPS}}\rangle = \sum_{\{\alpha\}} \mathbf{A}^{\alpha_1}\mathbf{A}^{\alpha_2}\cdots\mathbf{W}^{\alpha_i\alpha_{i+1}}\cdots\mathbf{A}^{\alpha_n}|\alpha_1\alpha_2\cdots\alpha_n\rangle,$$

3. Calculation of CAS amplitudes

$$T_1^{\mathrm{CAS}} = C_1 \qquad\qquad T_2^{\mathrm{CAS}} = C_2 - \frac{1}{2}(C_1)^2$$

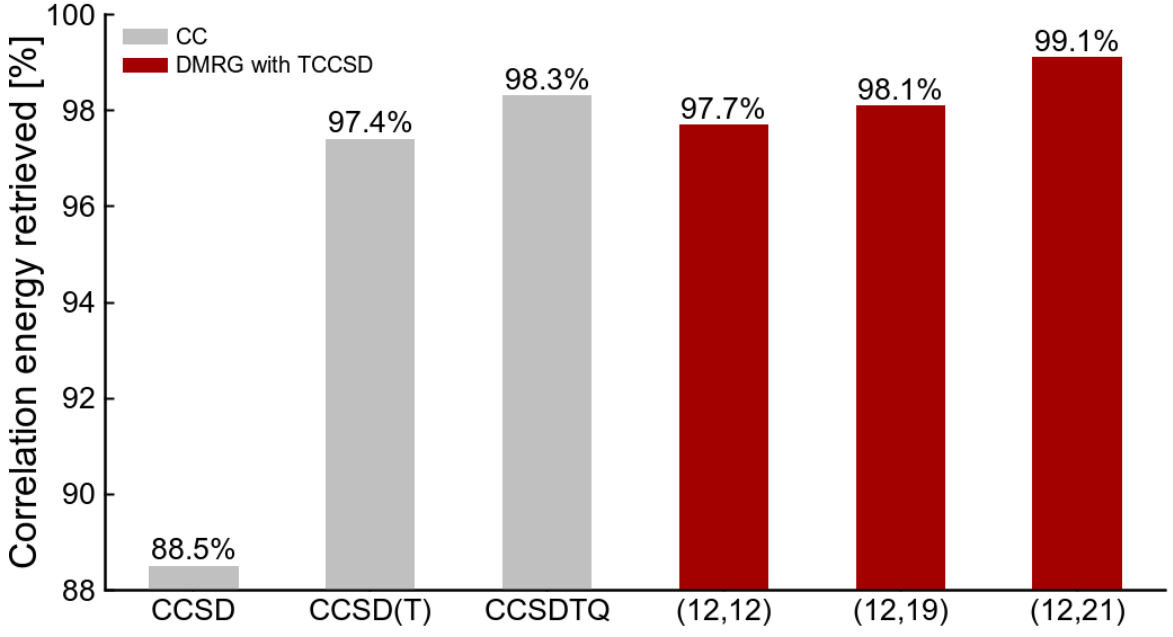
4. CCSD calculation for T_1^{ext} and T_2^{ext}



- Cost of the MPS $\rightarrow$ T_{12} conversion: $\mathcal{O}(M^2n^4)$ with a small prefactor, using techniques from Zgid and Nooijen, JCP **128**, 144115 (2008)

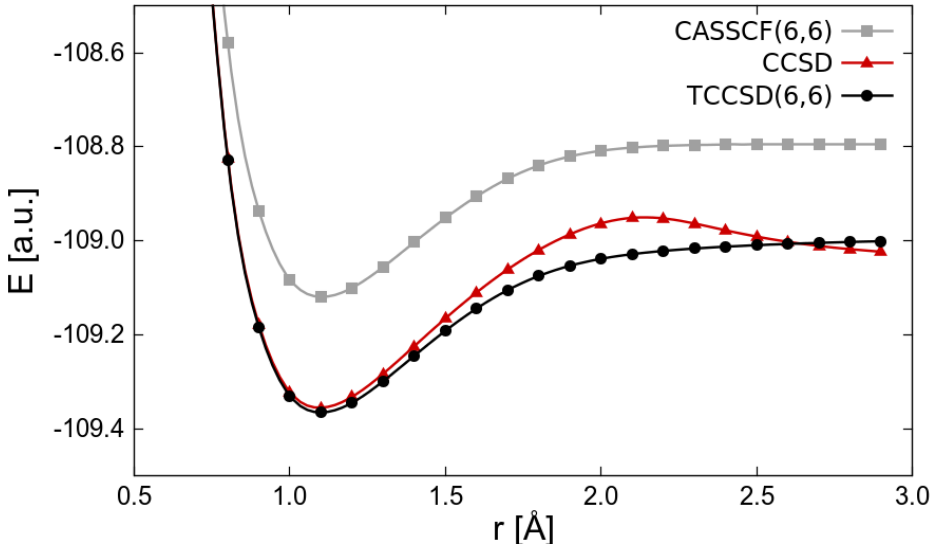
Chromium dimer – correlation energies

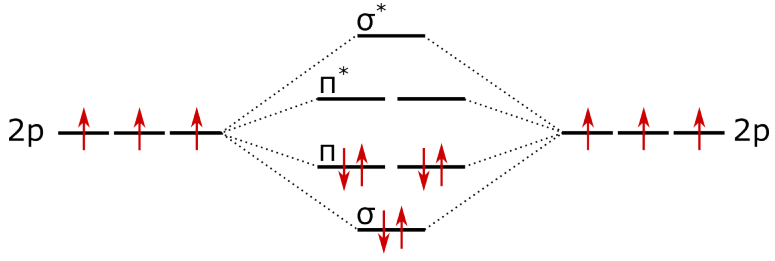
- Single-point calculation at 1.5 Å
- One-particle basis: RHF with Ahlrichs' SV basis set $\rightarrow$ (48e,42o)
- DMRG space selected based on $S^{(1)}$ profile
- DMRG performed with DBSS ($\epsilon_{\mathrm{tr}} \approx 10^{-7}$)
- Extrapolated DMRG by Olivares-Amaya et al. JCP 142, 034102, 2015 serves as a FCI benchmark



Nitrogen dimer

- Triple bond breaking resulting in six times degenerate HOMO
- CCSD yields unphysical barrier in PES





Mathematical analysis of the Tailored Coupled Cluster (TCC)

TCC approach was introduced as an alternative to the expensive and "knotty" multi-reference CC methods (MRCC). The TCC method divides the cluster operator into a **complete active space** (CAS) part, $\hat{S}$, and an **external space** (ext) part $\hat{T}$, i.e.,

$$|\Psi_{\mathrm{TCC}}\rangle = \exp(\hat{T})\exp(\hat{S})|\Psi_{\mathrm{HF}}\rangle \ .$$

Hence $\hat{T}$ and $\hat{S}$ commute, unlike in MRCC. The "linked" CC equations are now given by

$$\begin{cases} E = \langle\Psi_{\mathrm{HF}}|e^{-\hat{S}}e^{-\hat{T}}\hat{H}e^{\hat{T}}e^{\hat{S}}|\Psi_{\mathrm{HF}}\rangle \\ 0 = \langle\Psi_{\mu}|e^{-\hat{S}}e^{-\hat{T}}\hat{H}e^{\hat{T}}e^{\hat{S}}|\Psi_{\mathrm{HF}}\rangle \quad \text{for all} \quad \Psi_{\mu} \in \mathrm{CAS}^{\perp}. \end{cases} \tag{1}$$

- $|\Psi_{\mathrm{CAS}}\rangle = e^{\hat{S}}|\Psi_{\mathrm{HF}}\rangle$ is computed first and held fixed for the **dynamical correction step** by means of **CCSD** (CC with only single-double excitations) applied in $\mathrm{CAS}^{\perp}$.
- Although CC is **nonvariational**, it is **size-extensive** $\Longrightarrow$ inherited by TCC.
- **Important:** in TCC, the CAS cluster amplitudes are independent from the external space amplitudes, i.e. the TCC approach does not take coupling from the external space to the CAS into account!

The choice of the CAS using Quantum Information Theory

- **Notations:** N the number of electrons, K the number of spin-orbitals, k the "basis splitting number" ($N \leq k \leq K$) and $\mathcal{B} = \{\underbrace{\chi_1, \ldots, \chi_k}_{\mathcal{B}_{\text{CAS}}}, \underbrace{\chi_{k+1}, \ldots, \chi_K}_{\mathcal{B}_{\text{ext}}}\}$ the FCI basis.
- **Special cases:** $k = N \Longrightarrow$ CC (bad for static), and $k = K \Longrightarrow$ DMRG (bad for dynamic).
  Is there an optimal choice of k and error minimum in between?
- We choose the CAS space based the on the classification of the spin-orbital correlations dictated by the **mutual information** (a.k.a. **two-particle correlation**)

$$I(i,j) = S(\rho_{\{i\}}) + S(\rho_{\{j\}}) - S(\rho_{\{i,j\}}),$$

  where $S(\rho) = -\operatorname{Tr} \rho \ln \rho$ is the von Neumann entropy and $\rho_{\{\chi\}}$ is the reduced density matrix. Basis-dependent!
- **More precisely:** a pair (χ_i, χ_j) of spin-orbitals with...
  - **large** $I(i,j)$ are classified as **strongly correlated**,
  - **small** $I(i,j)$ are classified as **dynamically correlated**.
- This mutual information profile is obtained from a **quick**, low tensor rank DMRG calculation performed on the full system as a preliminary step.

Local analysis of the DMRG-TCC method

- The formulation (1) may be viewed as a **nonlinear Galerkin scheme**, i.e. it fits into a very wide class of numerical methods, which has a general mathematical framework.
- We have was shown that under certain assumptions the DMRG-TCC method admits a **locally unique** and **quasioptimal** solution.
  - **Local uniqueness:** the nonlinear equations (1) admit a unique approximate solution near the exact solution for a fixed CAS solution.
  - **Quasioptimality:** the "Galerkin solution" has the minimal error from the exact solution for a fixed basis set up to a multiplicative constant – a common feature of Galerkin-type methods.
- Instead of the conventional HOMO-LUMO gap, our **key assumption** is that there is a positive **CAS-ext gap** in the eigenvalues of the Fock operator, i.e. that $\boxed{\lambda_{k+1} > \lambda_k}$.

Error bounds: DMRG-TCC has a quadratic error bound

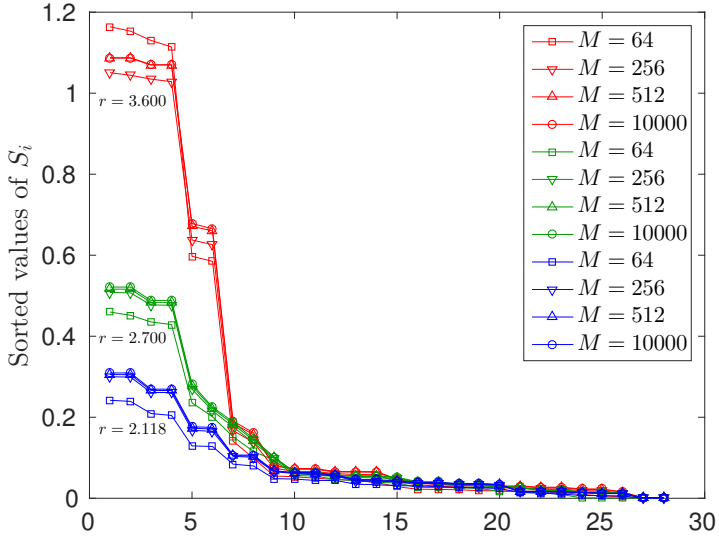
- The energy error ΔE of the DMRG-TCC method is measured from exact Full CI energy E, i.e. $H|\Psi^*\rangle = E|\Psi^*\rangle$.
- The **error bound** is given as $\Delta E \leq \Delta\varepsilon + \Delta\varepsilon_{\mathrm{CAS}} + \Delta\varepsilon^*_{\mathrm{CAS}}$ where
  - $\Delta\varepsilon$ **measures the truncation error introduced by restricting the CC method** to single-, and double excitations in the CCSD step, "tailored" by the DMRG solution on the CAS.
  - $\Delta\varepsilon_{\mathrm{CAS}}$ **measures the error of approximating the FCI solution with DMRG on CAS**, while the external part of the solution is held fixed; this in turn can be bounded by

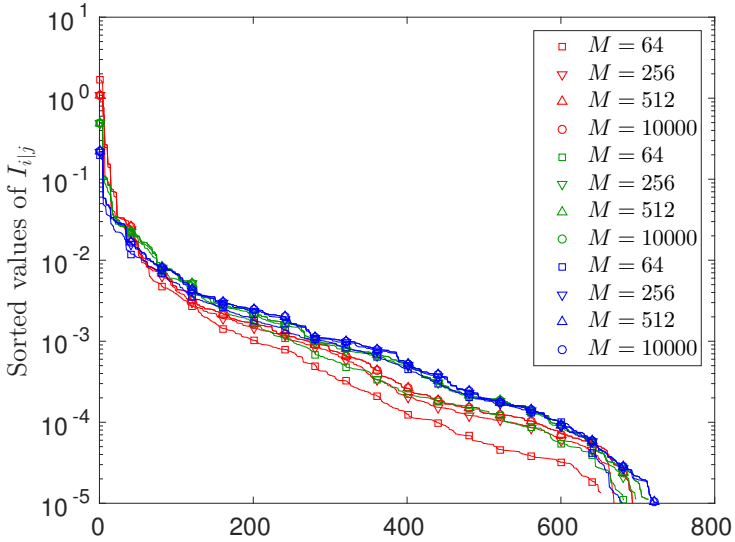
$$\Delta\epsilon_{\mathrm{DMRG}} \leq \Delta E_{\mathrm{DMRG}} + \|t_{\mathrm{CC}} - t^*_{\mathrm{CC}}\|^2_{\mathrm{ext}} + \|(\hat{S}_{\mathrm{DMRG}} - \hat{S}_{\mathrm{FCI}})\phi_0\|^2 + \sum_{|\mu|=1} \varepsilon_\mu (t^*_{\mathrm{CC}})^2_\mu,$$

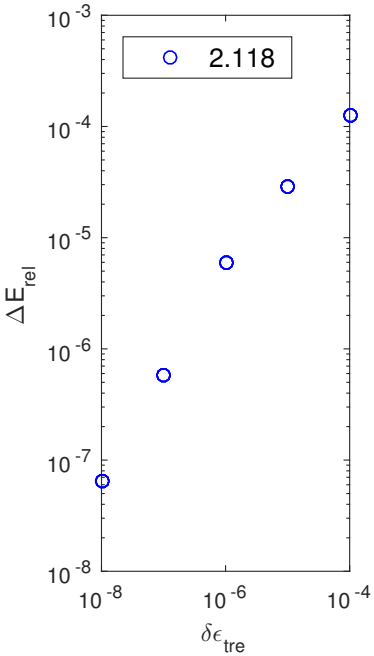
    where t_{CC} and t^*_{CC} are the approximate-, and exact cluster amplitudes, $\hat{S}_{\mathrm{DMRG}}$ and $\hat{S}_{\mathrm{FCI}}$ are the cluster operators and $\epsilon_\mu = \epsilon^{A,\ldots,A_n}_{I_1,\ldots,I_n} = \sum_{j=1}^{n}(\lambda_{A_j} - \lambda_{I_j})$, with λ_j denoting the eigenvalues of the Fock operator. Here, ΔE_{DMRG} can be made arbitrarily small. The last term is a **"methodological error"** inherent in the TCC method, which is small in applications.
  - $\Delta\varepsilon^*_{\mathrm{CAS}}$ **measures the error between the full exact solution and solution obtained by FCI on CAS and untruncated CC.**

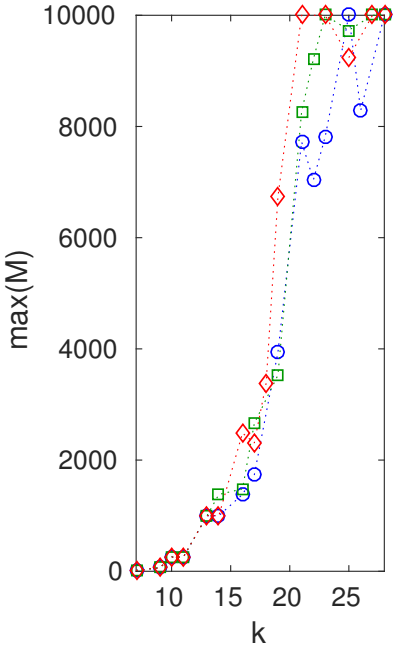
Numerical error analysis on the N_2 ($N = 14$ e, $K = 28$ orb)

DMRG for the full orbital space, CAS is formed from $k = K = 28$ orbitals



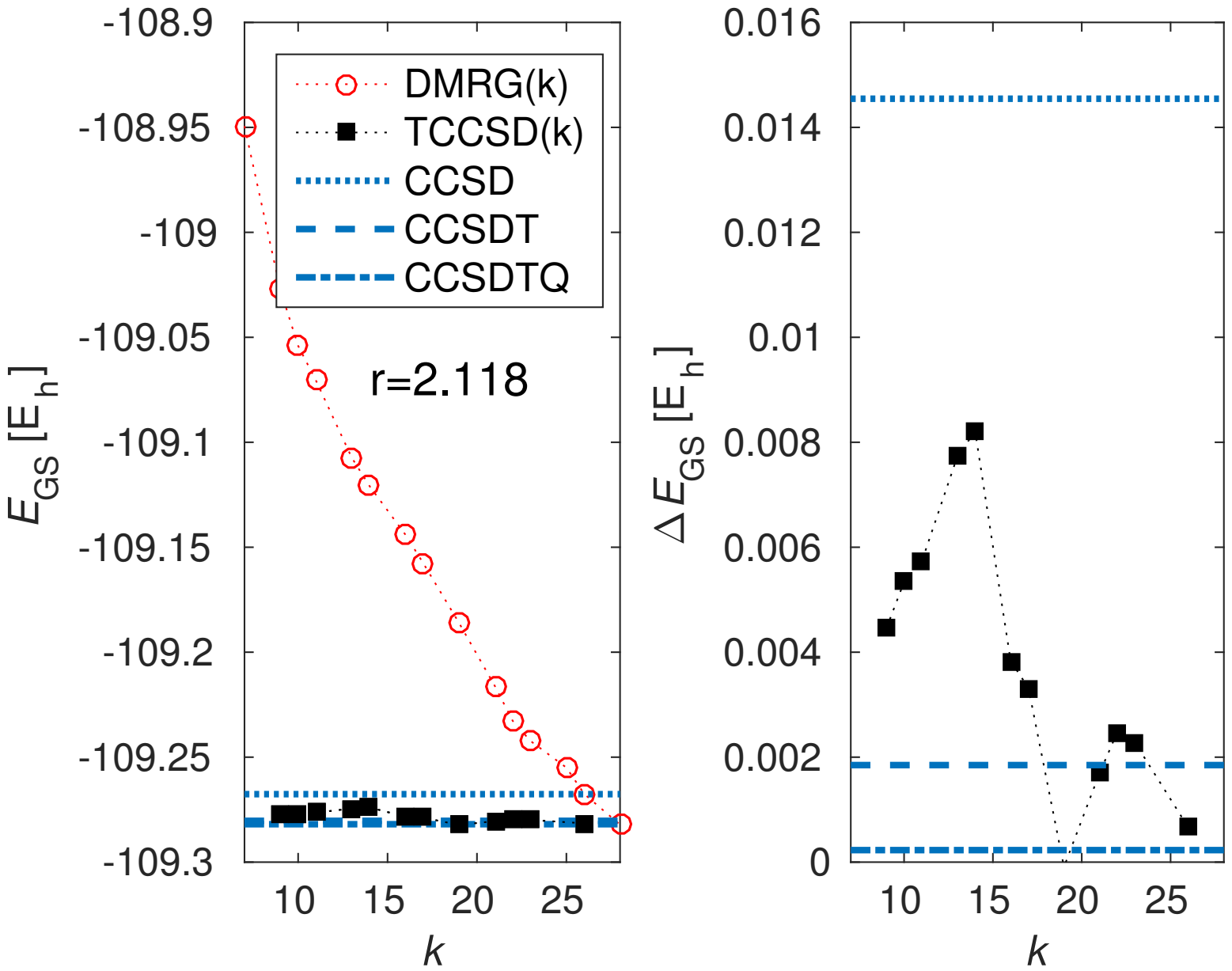




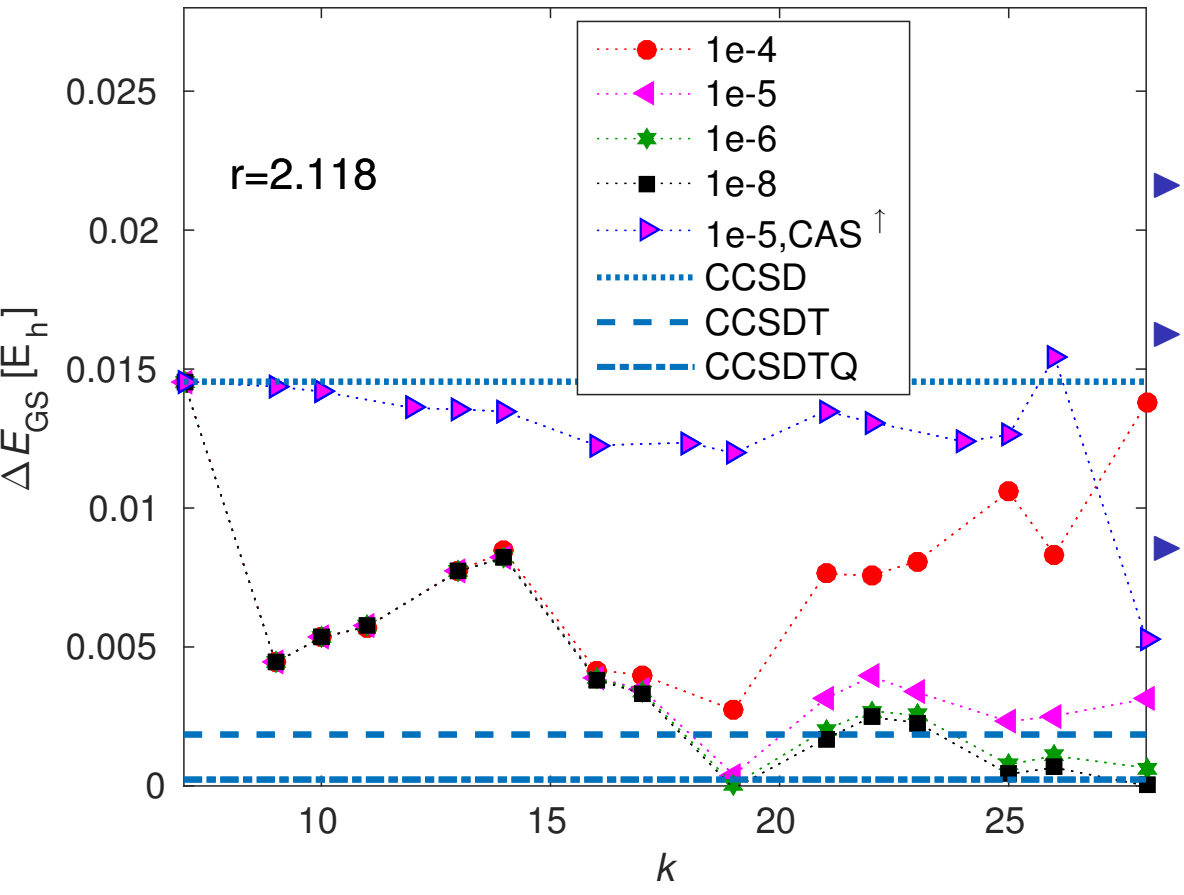


- r=2.118 a_0, 2.700 a_0, 3.600 a_0
- S_i shifts upward
- I_{ij} exponential tail not effected
- static and dynamic correlations
- extrapolation with $\delta\varepsilon_{\mathrm{Tr}}$
- E_{FCI} = CCSDTQPH
- CAS-vector

$N/2 \leq k \leq K$ dependence at equilibrium geometry



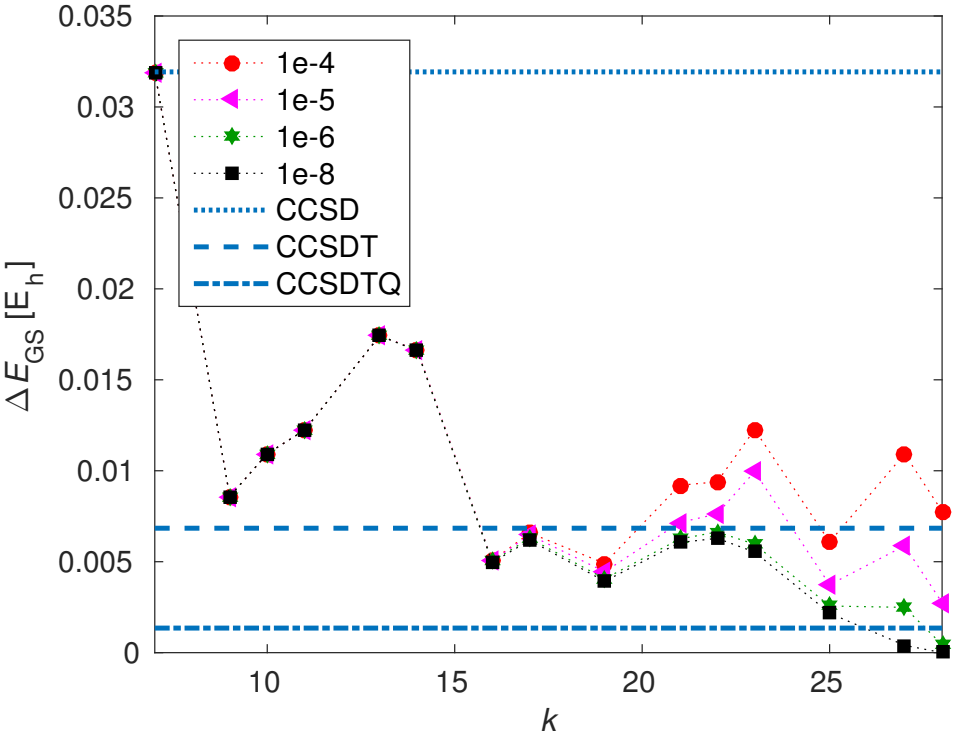
$N/2 \leq k \leq K$ dependence at equilibrium geometry

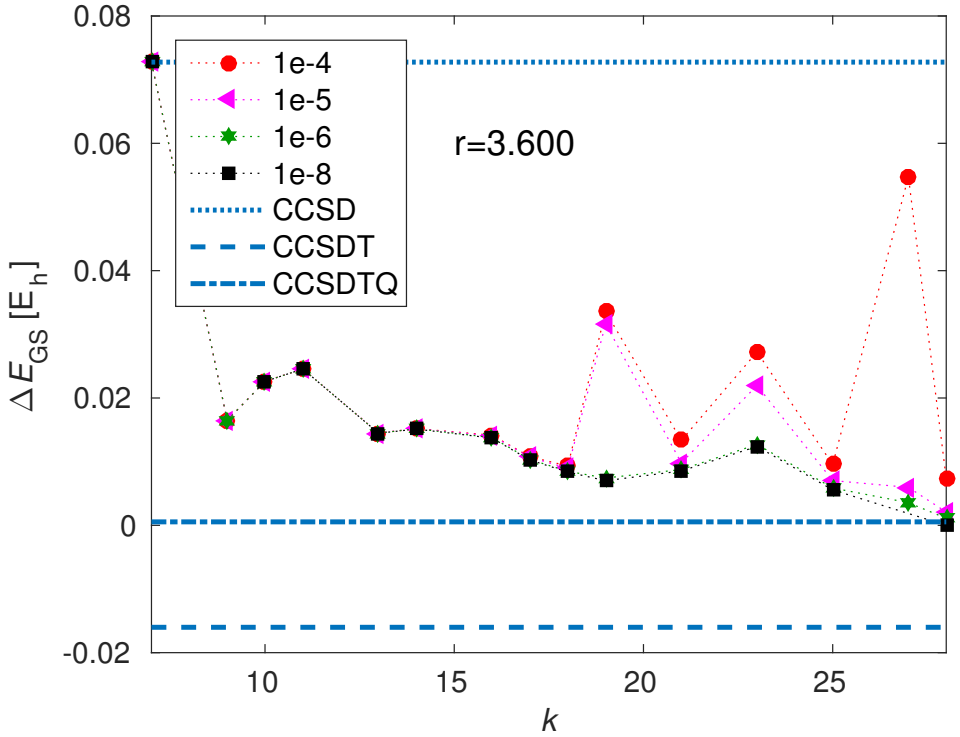


- CAS formed according to CAS-vector
- DMRG-TCCSD energy is far below the CCSD energy for all CAS, even for $k = 9$
- irregular error behavior, small values for various k-s (methodological error)

- supports k-dependent constant in the mathematical analysis
- determine optimal k value from the computational point of view
- effect of truncation error and CAS choice (CAS$^{\uparrow}$)

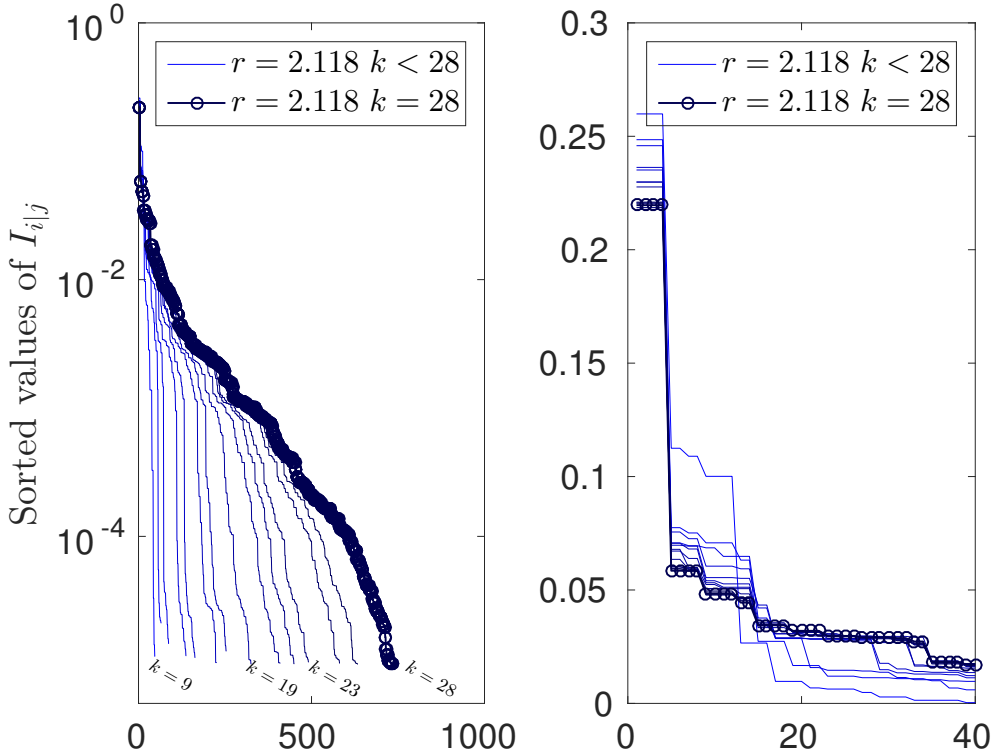
$N/2 \leq k \leq K$ dependence for stretched geometries

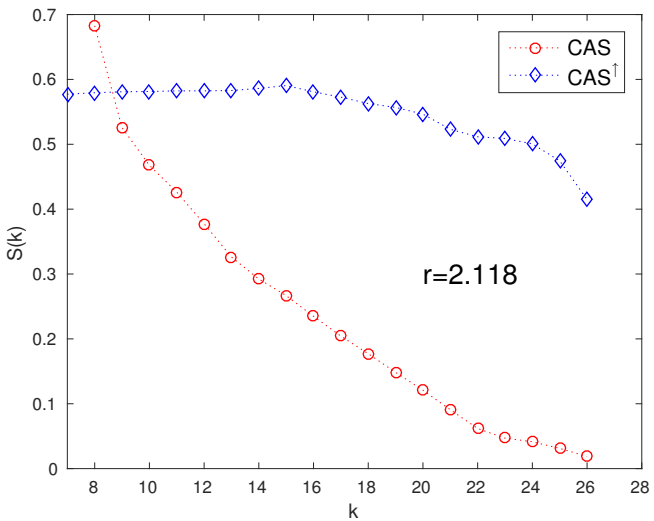




- Multi-reference character of the wave function is more pronounced
- This becomes apparent through the entropy profiles
- For $r = 3.600\, a_0$ the CC computation fluctuates with increasing excitation ranks and CCSDT is even far below the FCI reference energy, revealing the variational breakdown of the single-reference CC method for multi-reference problems.
- DMRG-TCCSD is stable along the whole PES!

Entropy Error Analysis





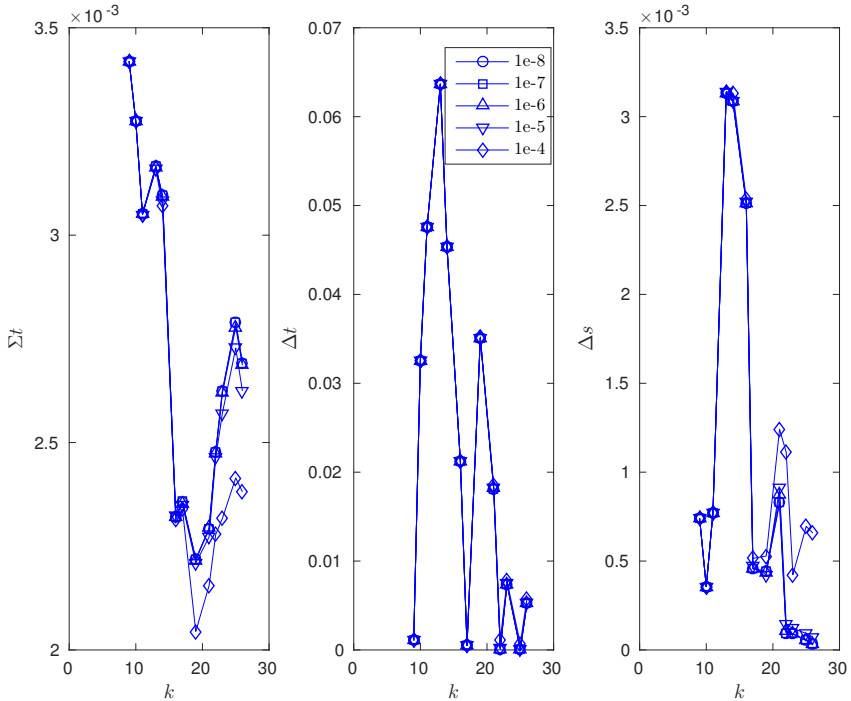
- largest values of I_{ij} change only slightly with incensing k
- exponential tail of $I_{i,j}$ becomes more visible for larger k
- CAS-Ext correlations can also be simulated by a DMRG

- block entropy $S(\rho_{\mathrm{CAS}(k)})$ as a function of k
- block entropy decays monotonically cannot explain irregular error profile

Amplitude Error Analysis

$$
\begin{aligned}
e(k,\delta_{\varepsilon_{\mathrm{Tr}}}) &= \sum_{\substack{\mu:\\ \mu=1}} \left(t_{\mathrm{CCSD}}(k,\delta_{\varepsilon_{\mathrm{Tr}}})\right)_{\mu}^{2} \\
&+ \sum_{\substack{\mu:\\ \mu=1,2}} \left[\left(t_k^* - t_{\mathrm{CCSD}}(k,\delta_{\varepsilon_{\mathrm{Tr}}})\right)_{\mu}^{2} + \left(s_k^* - s_{\mathrm{DMRG}}(k,\delta_{\varepsilon_{\mathrm{Tr}}})\right)_{\mu}^{2}\right].
\end{aligned}
$$

Here the valid index-pairs are $\mu = (\mathbf{i}, \mathbf{a})$, with $\mathbf{i} = (i_1, \ldots, i_n) \in 1, \ldots, N/2^n$, and $\mathbf{a} = (a_1, \ldots, a_n) \in N/2+1, \ldots, K^n$. The excitation rank is given by $\mu = n$ where $n = 1$ stands for singles, $n = 2$ for doubles, and so on.



Conclusion

- T3NS is a very challenging new tensor format.
- TNS-TCCSD is very efficient method to recover both static and dynamic correlations.
- DMRG-TCCSD is size-extensive, admits a locally unique and quasioptimal solution.
- DMRG-TCCSD has a quadratic error bound
- As demonstrated on the N_2 molecule DMRG-TCCSD is stable along the whole PES! $\rightarrow$ black-box implementation possible.
- Other extensions of TCC with LPNO, etc , excited states, ...
- Need further mathematical analysis

Future: Migration of DMRG/TNS into the NWCHEM professional software package based on its massively parallelized tensor library (Pacific North National Laboratory) Supports: Lendület grant, Hungarian Academy of Sciences, the Hungarian National Research, Development and Innovation Office (K120569), Hungarian Quantum Technology National Excellence Program (Project No. 2017-1.2.1-NKP-2017-00001), European Research Area(ERA, DFG, EU (SIQS, RAQUEL, AQUS) the BMBF, and the