Optimizing density-matrix renormalization group method using quantum information entropy

Ö. Legeza and J. Sólyom

Research Institute for Solid State Physics,

H-1525 Budapest,

P. O. Box 49, Hungary

† *In memoriam of B. A. Hess.*

Chair of Theoretical Chemistry,

Friedrich–Alexander University Erlangen–Nürnberg.

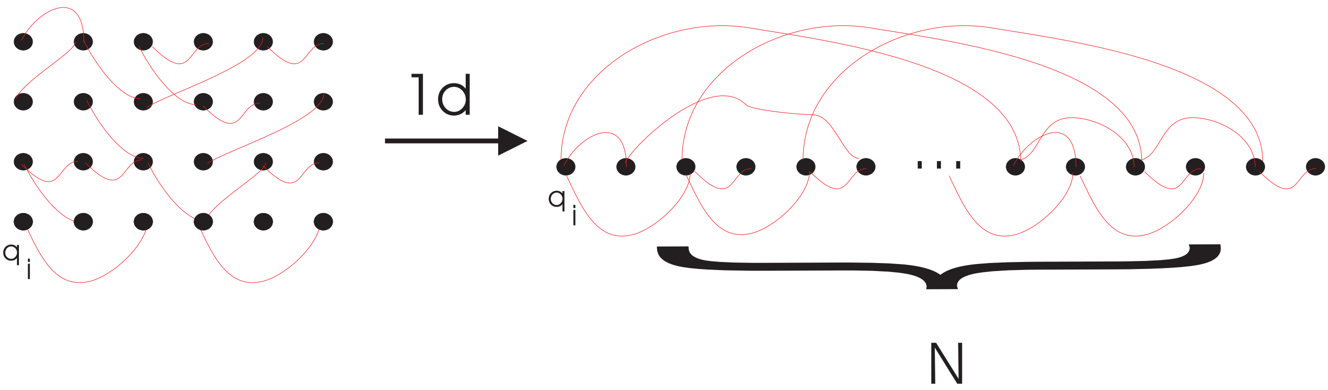
Chair of Theoretical Chemistry, Bonn.

(Dated: March 22, 2006)

I. TOPICS TO BE COVERED

- Model Hamiltonian
- Motivation to study quantum information theory
- Basics of Information Theory and connection to DMRG
- Separability of the target state, Schmidt decomposition
- von Neumann quantum information entropy
- Entanglement
- Quantum data compression
- Quantum information generation
- Entanglement localization vs. interaction localization
- Dynamically Extended Active Space (DEAS) procedure

II. INTERACTING QUANTUM SYSTEM AND MODEL HAMILTONIAN



- Each site represents a quantum object with q_i degrees of freedom
- Size of the Hilbert space (Λ) is: $\dim \Lambda = \Pi_i \, q_i$
- Operators are defined for each site, for example $S_i^\dagger, S_i^z$ for spin systems, $c_i^\dagger$, c_i for fermion systems

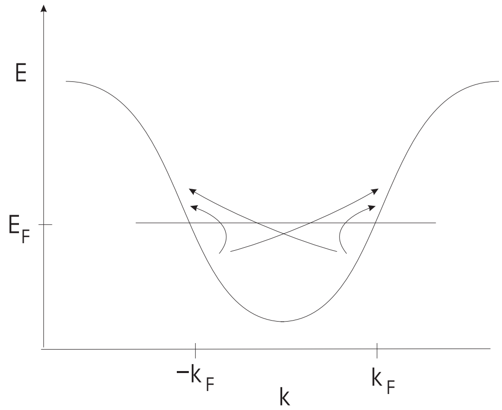
• The system is described by the Hamiltonian, for example,

$$\mathcal{H} = \sum_{ij\sigma} T_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + \sum_{ijkl\sigma\sigma'} V_{ijkl} c^{\dagger}_{i\sigma} c^{\dagger}_{j\sigma'} c_{k\sigma'} c_{l\sigma}, \tag{1}$$

where T_{ij} denotes the matrix elements of the one-particle Hamiltonian and V_{ijkl} stands for the matrix elements of the electron interaction operator.

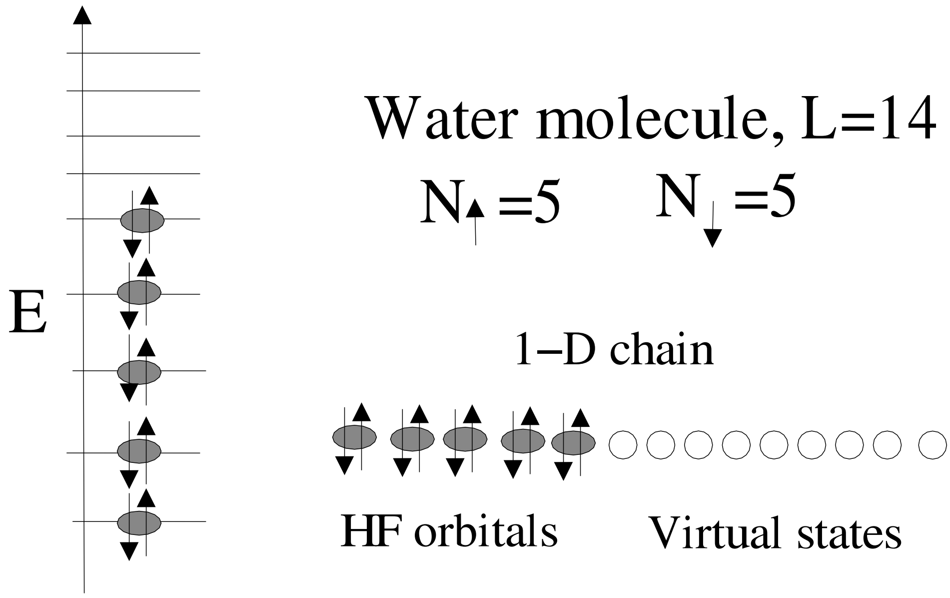
• Depending on the structure of T_{ij} and V_{ijkl} this Hamiltonian can describe a **molecule** or a **usual fermionic model** in solid state physics, e.g., the Hubbard or extended Hubbard models in one or higher dimensions or for example coupled fermion chains (T. Xiang, S.R. White).

- The indices i, j, k, l denote momenta with $k_i = (2\pi n)/L, -L/2 < n \leq L/2$.
- $T_{ij} = -2t\cos(k_i)\delta(i-j)$ and $V_{ijkl} = (U/L)\delta(i+j-k-l)$. U is given in units of t with $t = 1$.
- Dispersion relation. Arrows indicate the most probable electron-hole excitations with small momentum transfer in the small U limit.



IV. QUANTUM CHEMISTRY APPLICATION (QC-DMRG)

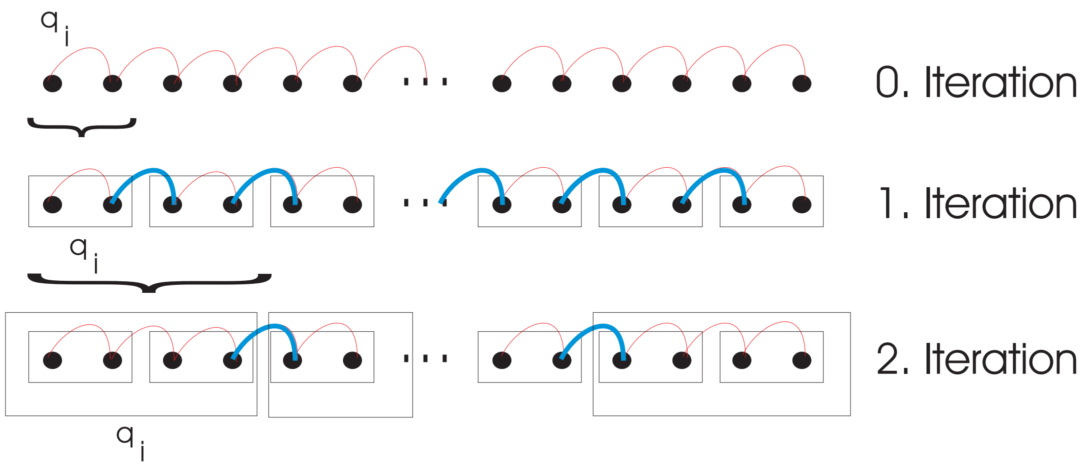
• In the case of quantum chemistry a one-dimensional chain is built up from the molecular orbitals that were obtained, e.g., in a suitable mean-field or MCSCF calculation (MOLPRO program package).



• Exact diagonalization → exact solution, (not possible for large systems)

• Renormalization group method: selecting out the relevant part of the energy spectrum to describe the physical properties of the system, i.e. integrating out high energy part of the energy spectrum

Block Renormalization



1) Forming and diagonalizing the Hamilton operator for two sites ($\dim \Lambda = q^2$)

2) Selecting q states out of the q^2 eigenstates (corresponding to lowest eigenvalues)

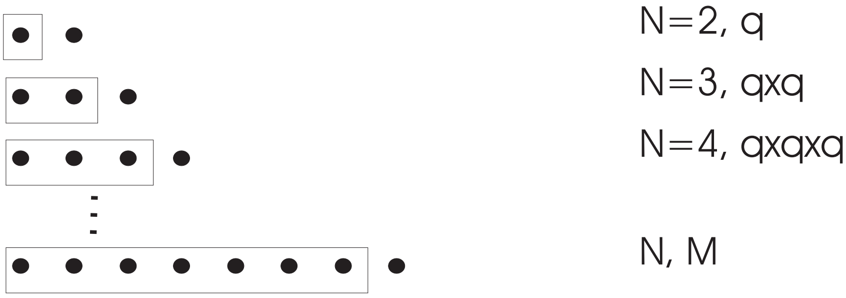
3) Forming a transformation matrix (O) from the q selected eigenstates

4) Renormalizing all site operators according to $A_{rmblock} = OAO^\dagger$ to form block representation

5) $\rightarrow$ step 1 or stopping if further application of renormalization gives no new result (fix-point)

- Site degrees of freedom and number of operators remain the same, scale equations, flow diagrams, fix-points

Wilson's numerical renormalization



1) Forming and diagonalizing the Hamilton operator for two sites

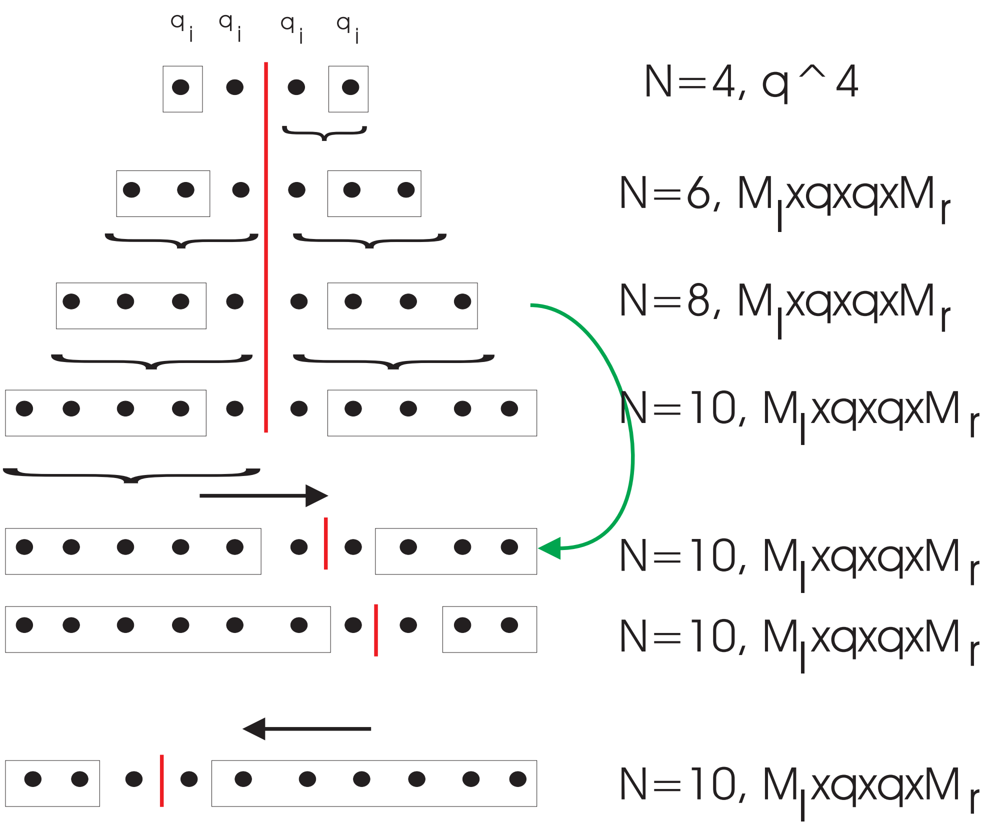
2) Selecting $M > q$ states out of the $M \times q$ eigenstates (corresponding to lowest eigenvalues)

3) Minimizing energy, scaling behavior of quantities as a function of length

• Site degrees of freedom and number of operators change, no scale equations, no flow diagrams, solution requires numerical approach, relevant, irrelevant marginally relevant operators

Density-matrix renormalization group method

Super-block configuration



1) Forming and diagonalizing the superblock Hamilton operator

$$|\Psi_{\mathrm{TG}}\rangle = \sum_{\alpha ij\beta} \psi_{\alpha ij\beta} |\phi_\alpha^{(l)}\rangle \otimes |\phi_i^{s_l}\rangle \otimes |\phi_j^{s_r}\rangle \otimes |\phi_\beta^{(r)}\rangle \qquad (2)$$

where $\psi_{\alpha ij\beta}$ coefficients are determined by the diagonalization of the superblock Hamiltonian.

2) Forming bi-partite representation $|\Psi_{\mathrm{TG}}\rangle = \Sigma_{ij}\, \psi_{i,j} |\phi_i^{(L)}\rangle |\phi_j^{(R)}\rangle$

3) Forming reduced subsystem density matrix $\rho_{i,i'} = \Sigma_j\, \psi_{i,j} \psi^\dagger_{i',j}$

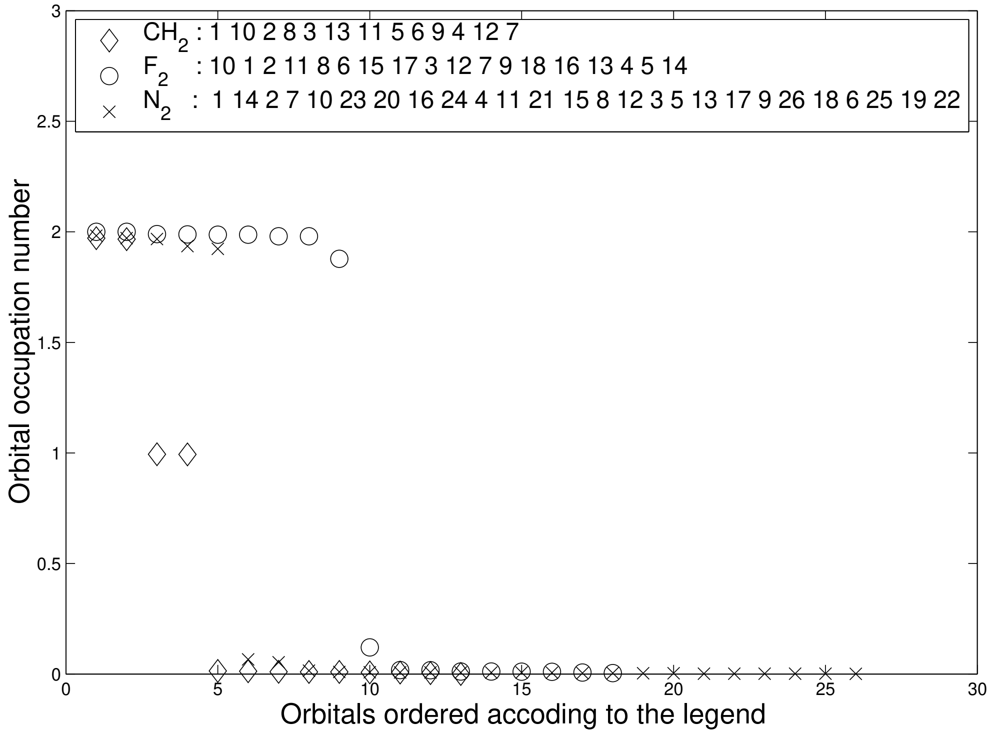
4) Diagonalizing $\rho \to \omega_\alpha$ eigenvalues, $|\phi_\alpha\rangle$ eigenstates

5) Forming O matrix using M selected $|\phi_\alpha\rangle$ eigenstates corresponding to the M largest ω_α

6) Truncation error: $E_{TR} = 1 - \Sigma_{\alpha=1}^{M}\, \omega_\alpha$

•

• Decay of orbital occupation number obtained in the full CI limit and the corresponding ordering of orbitals for a few selected molecules.



V. TECHNICAL DIFFICULTIES DUE TO THE LONG RANGE INTERACTIONS

• One needs auxiliary operators containing various combinations of renormalized operators.

• In the general case the DMRG superblock Hamilton is decomposed as (Xiang):

$A_L^{(4)}$	1 term
$A_R^{(4)}$	1 term
$A_L^{(3)}\ A_R^{(1)}$	2× R terms
$A_L^{(1)}\ A_R^{(3)}$	2× L terms
$A_L^{(2)}\ A_R^{(2)}$	$6 \times R^2$ terms

VI. MOTIVATION TO APPLY QUANTUM INFORMATION ENTROPY

• For homogeneous lattice models studied by standard real space DMRG with periodic boundary condition each lattice site is equivalent and carries the same amount of information.

• In contrast to this, k-points or molecular orbitals lying closer to or farther away from the Fermi surface have different information content.

• The new method is very sensitive to the ordering of the k-points or molecular orbitals and an optimal ordering would have a major impact on the performance of MS-DMRG.

VII. OUR RESULTS BY APPLICATION OF QUANTUM INFORMATION ENTROPY

- New initialization procedure for MS-DMRG and QC-DMRG that extends the active space in a dynamic fashion (DEAS procedure). It provides 10^{-3}-10^{-4} accuracy after the first half sweep and reduces computational time by a factor of 10-100.

- Optimization of ordering based on the competition between entanglement localization and interaction localization.

- Optimization of computational resources by reducing number of block states using results of quantum data compression.

- A new truncation procedure based on Kholevo's accessible information: modified DBSS procedure.

• A sum rule which relates sites entropies and the total information generated by the RG procedure.

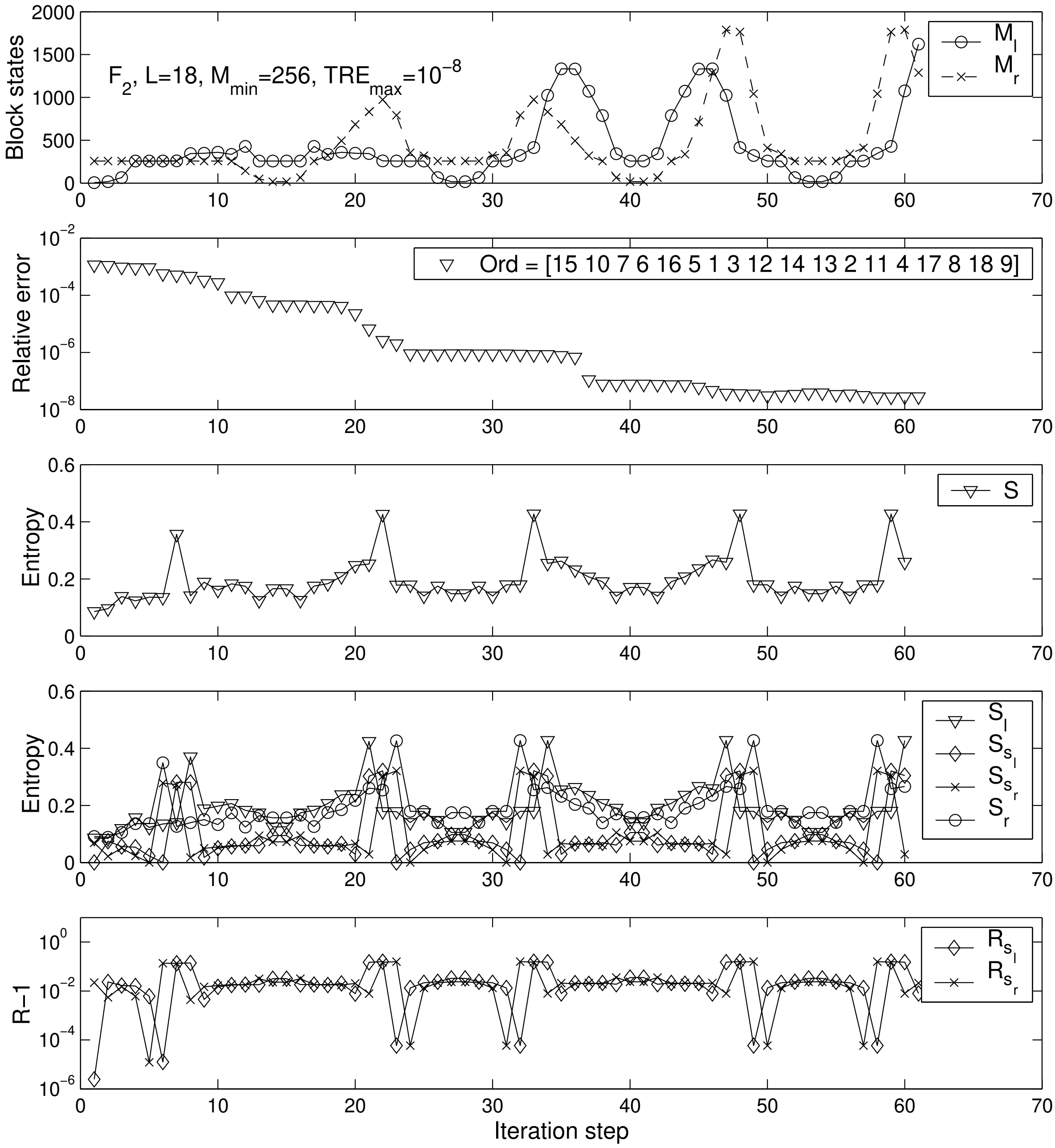
Benchmark results:

• E_{GS} of F_2 molecule (18/18) up to 8 d.p. using present day FCI QC program on an Origin Memory Server (Erlangen) takes 50-60 hours using 24-30GB of RAM.

$\Delta E_{\mathrm{GS}} = 10^{-}6$ using QC-DMRG takes 20 minutes, 1GB RAM on a PIV-2GHz Notebook.

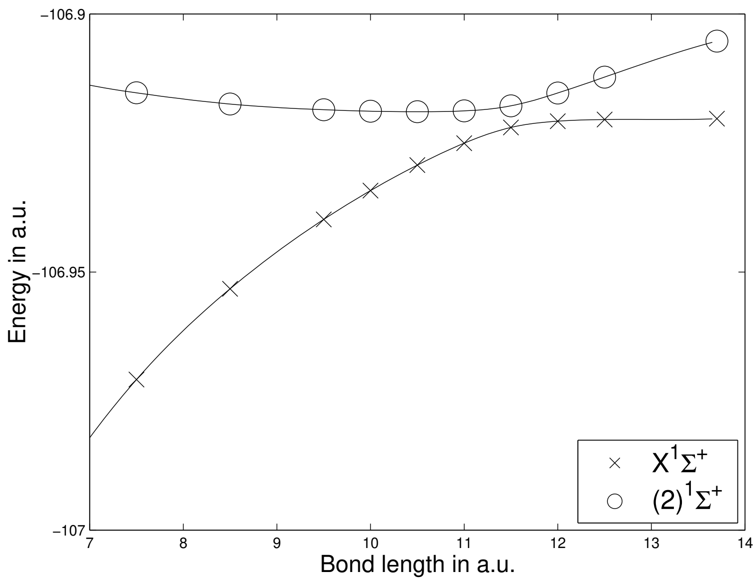
• The ground state energy of H_2O (10/41) with QIT optimized DMRG up to 4d.p. : 1cpu, (PIV-2GHz Notebook), 1GB RAM, takes 4h

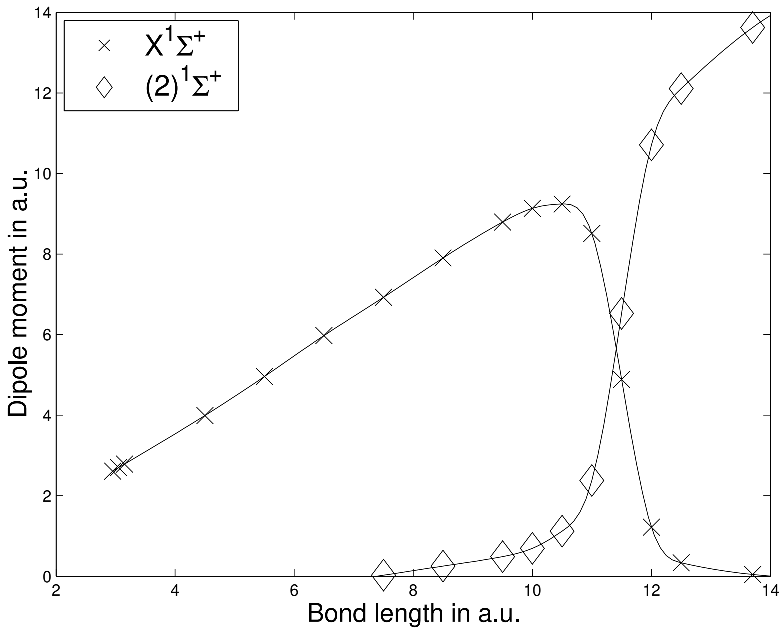
• In the DBSS approach the truncation error is kept fixed and the number of retained states is varied dynamically, which allows to control the accuracy more precisely. A threshold value on the minimum number of block states M_{min} has also been introduced in order to avoid higher lying local minima. Example: F_2 (18/18).



• Present stage: Correlating 30-40 electrons on 60-100 orbitals. Hopefully this is not the end ...

• QC-DMRG can treat easily those problems for which standard QC approaches fail. Example 1: Ionic–neutral curve crossing of LiF





- Dipole moment is a continuous function of the bond length.
- Major aim is to get faster performance than Couple Cluster methods.

VIII. BASICS OF INFORMATION THEORY AND CONNECTION TO DMRG

- What is information

- We all carry information (different properties)

- Need interaction (correlation) to “measure” it

- Need selection rules to determine which properties are relevant and which are irrelevant

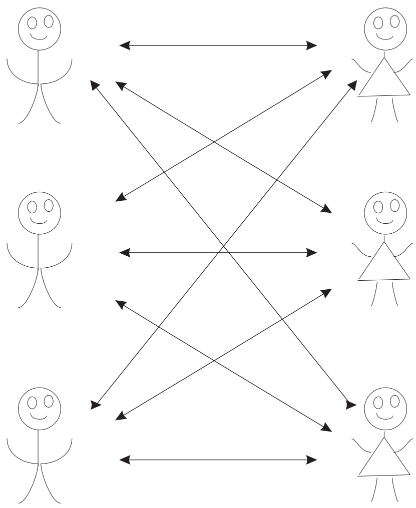
- In a dynamic system properties can be dominant or recessive

- In case of DMRG the total information of the system is shared between two blocks, thus blocks also carry information

- In case of DMRG blocks evolve as a function of iteration steps

- Need a protocol to measure information content of the blocks

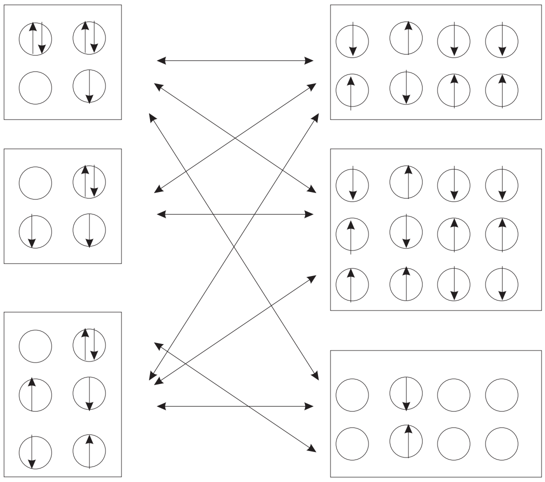
- Example 1: “Real life”



- Properties: nice, smart, tall, color of hair, color of eyes, etc.
- Need interaction to exchange information
- Need selection rules to weight properties

• Example 2: Quantum chain with N=6 sites and for a partitioning of L=2, R=4

• q=4 basis states: $|0\rangle$, $|\downarrow\rangle$, $|\uparrow\rangle$, $|\downarrow\uparrow\rangle$



• "Individuals" $\equiv$ blocks with different subset of the q^l and q^r states

• Total information of the system is shared between the two blocks

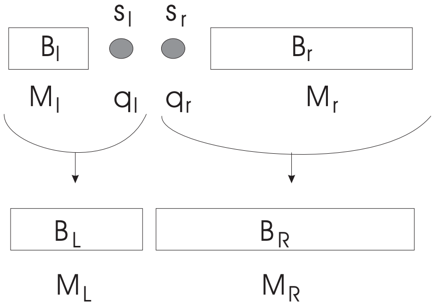
• Need selection rules to weight importance of the basis states,

• For a given target state only a subset of the left and right block states have non-zero weight

• Meaning to a message can be assigned iff response of receiver is taken into account: (coding, decoding)

• Information content of both blocks depend on each other $\rightarrow$ **mutual information**

IX. DMRG IS A COMPOSITE SYSTEM



• The so called target state ($|\Psi_T\rangle$) of the superblock system which is the eigenstate that one wants to calculate is formed from the direct product states of the blocks and the sites.

• If the target state is a pure state it follows from the Schmidt decomposition that

$$|\Psi_T\rangle = \sum_{i=1}^{r\leq\min(M_L,M_R)} \omega_i |e_i\rangle \otimes |f_i\rangle, \tag{3}$$

where $|e_i\rangle \otimes |f_i\rangle$ form a bi-orthogonal basis and $\Sigma_i \, \omega_i^2 = 1$.

• Even if the target state is a pure state, the left and right blocks are in a mixed state described by the density matrices $\rho_L = \mathrm{Tr}_R \rho$, and $\rho_R = \mathrm{Tr}_L \rho$, respectively.

• Analogously one can define

$\rho_l = \mathrm{Tr}_{s_l} \mathrm{Tr}_R \rho$,

$\rho_r = \mathrm{Tr}_L \mathrm{Tr}_{s_r} \rho$,

$\rho_{s_l} = \mathrm{Tr}_l \mathrm{Tr}_R \rho$,

$\rho_{s_r} = \mathrm{Tr}_L \mathrm{Tr}_r \rho$.

• A quantitative characterization of the degree of mixtures is given by the von Neumann entropy. For of a finite system of length N, with density matrix $\rho^{(N)}$

$$S(\rho^{(N)}) = -\mathrm{Tr}(\rho^{(N)} \ln \rho^{(N)}) \,, \tag{4}$$

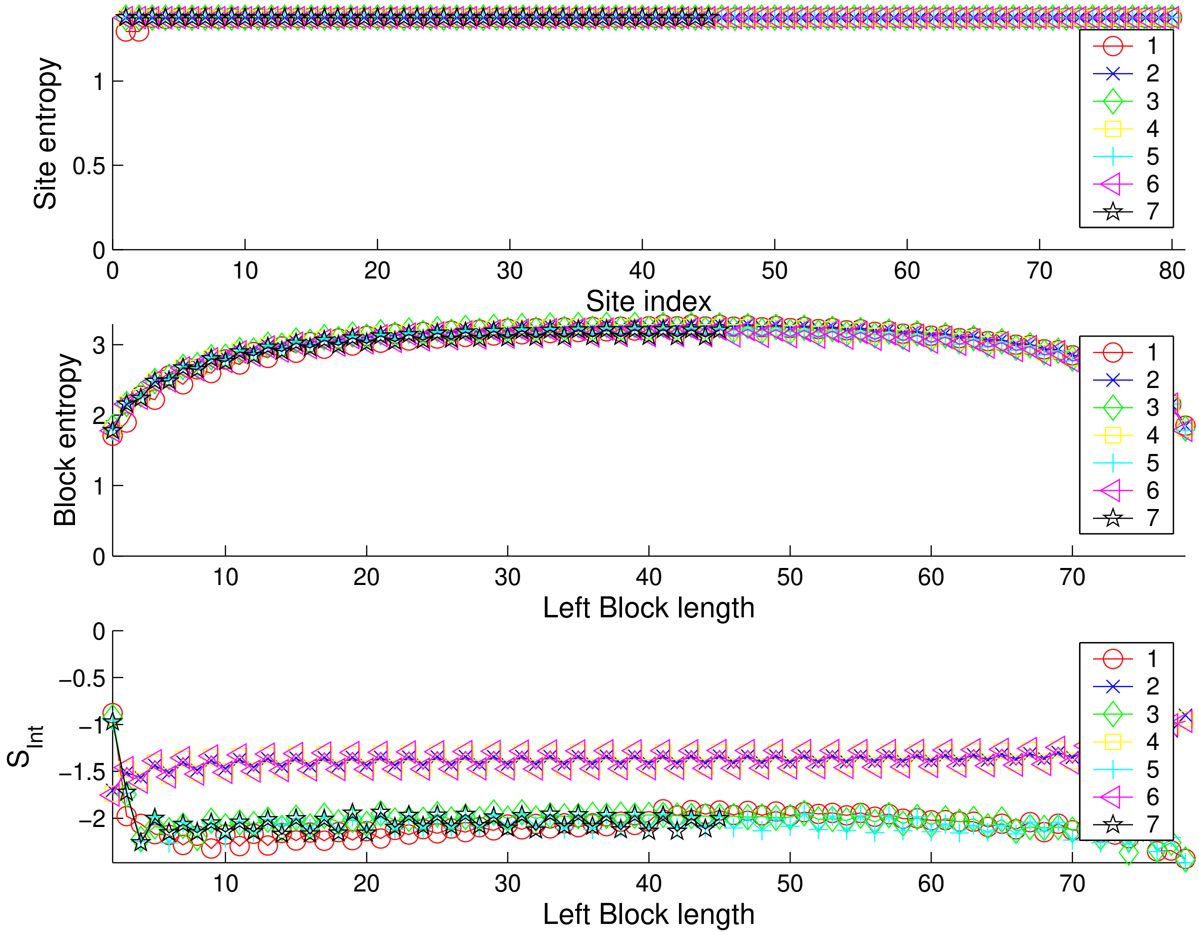
• It is zero for a pure state and $S = \ln N$ for a totally mixed states with $\rho = \frac{1}{N}\mathbb{1}$.

• For a pure target state S is also related to the entanglement

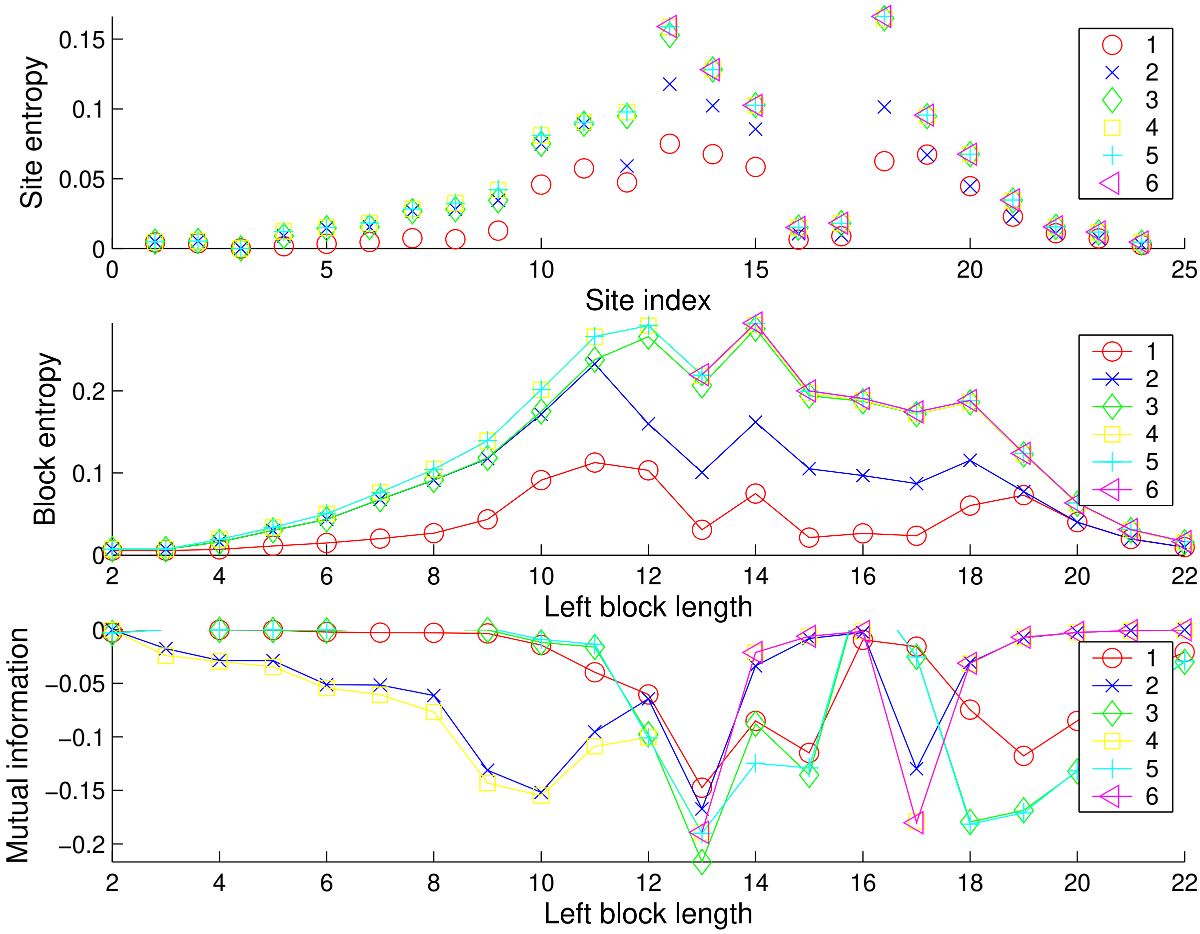
• G. Vidal *et al*: S diverges logarithmically with the system size if the system is critical and the spectrum is gapless, while it saturates as N increases for non-critical, gapped models.

• Site entropy: $S(\rho_{sr})$

• Ex.1: Half filled 1-d Hubbard chain, $M = 512$, N=80, PBC



- Example 2: H_2O molecule (8/24) with $M_l = M_r = 64$



X. QUANTUM DATA COMPRESSION

• **Pure-state ensemble:** Message consisting N letters $\rho^{(N)} = \rho \otimes \rho \otimes ... \otimes \rho$, where $\rho = \Sigma_\alpha \, p_\alpha |\phi_\alpha\rangle\langle\phi_\alpha|$, and $S(\rho^{(N)}) = NS(\rho)$.

• Hilbert space of the entire message, $\Lambda^{(N)} = \Lambda \otimes \Lambda \otimes \cdots \Lambda$

• A projection $\Pi_{\rm typ}$ and $\Pi_{\rm atyp}$

• If for any $\epsilon > 0$ almost all weight of the ensemble lies within $\Lambda^{(N)}_{\rm typ}$ and $Tr\Pi_{\rm typ}\rho^{(N)}\Pi_{\rm typ} > 1-\epsilon$ then for any δ and $N \to \infty$

$$e^{-N[S(\rho)+\delta]} < \omega_n < e^{-N[S(\rho)-\delta]} \, , \tag{5}$$

$$(1-\epsilon)e^{N[S(\rho)-\delta]} \leq \dim \Lambda^{(N)}_{\rm typ} \leq e^{N[S(\rho)+\delta]} \, , \tag{6}$$

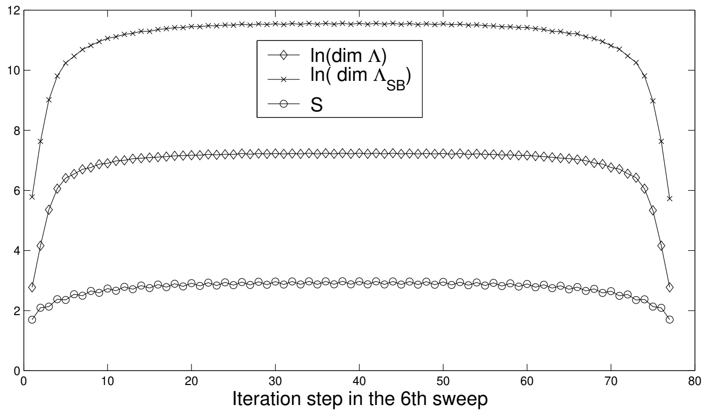
• In general for a quantum lattice $\rho^{(N)} \neq \rho \otimes \rho \otimes ... \otimes \rho$ and $S(\rho^{(N)}) \neq NS(\rho)$.

• The Hilbert space of the system (called superblock Hilbert space) is defined on a bi-partite system, $\Lambda_{\rm typ} = \Lambda^{(L)}_{\rm typ} \otimes \Lambda^{(R)}_{\rm typ}$.

• The $\Pi^L_{\rm typ}$ and $\Pi^R_{\rm typ}$ projection operators are formed from the M chosen eigenstates of ρ_L and ρ_R and due to the truncation of basis states ϵ is always greater than zero.

• In the DBSS approach the truncation error is kept fixed and the number of retained states is varied dynamically, which allows to control the accuracy more precisely. A threshold value on the minimum number of block states M_{min} has also been introduced in order to avoid higher lying local minima.

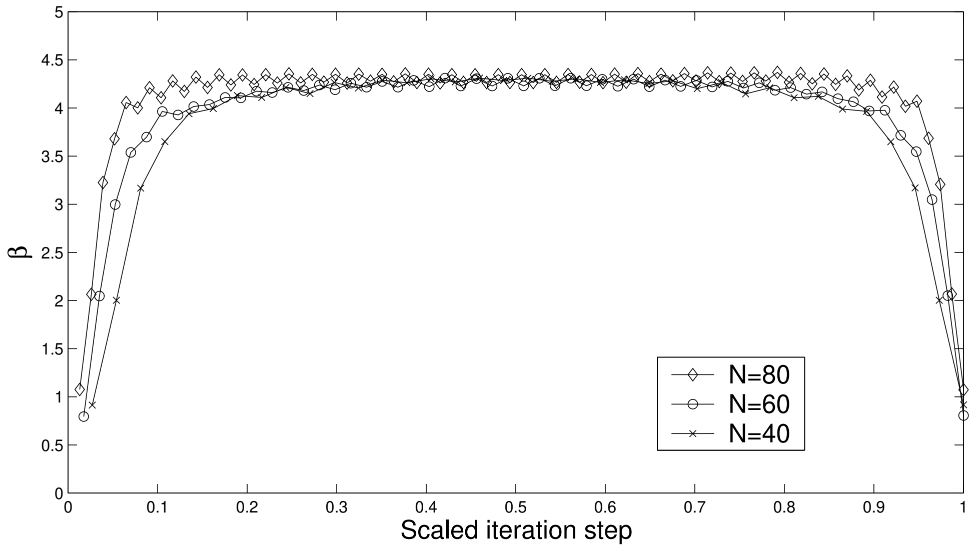
• Figure shows $\ln(\dim \Lambda)$, $\ln(\dim \Lambda_{SB})$ and the block entropy obtained after the 6th sweep for $U = 1$, $N = 80$ as a function of iteration step shifted modulo 80 using DBSS approach for $TRE_{max} = 10^{-4}$ and $M_{min} = 16$.



• A simple relationship exists between the two quantities:

$$\beta \equiv \ln(\dim \Lambda) - S \, . \tag{7}$$

• Figure shows the value of β for $U = 1$, $TRE_{max} = 10^{-4}$ for three different chain lengths, $N = 40, 60, 80$. β is plotted as a function of the shifted iteration step rescaled by the length of the chain.



• After some six sweeps except for very short block lengths β is practically independent of N for a given accuracy.

• **Mixed-states ensemble**: The mutual information between the sender's and receiver's measurement is bounded from above by

$$I \leq S(\rho) - \sum_a p_a S(\rho_a)\,, \tag{8}$$

where $\rho = \Sigma_a\, p_a \rho_a$ and S is the von Neumann entropy.

• The lower bound on accessible information (Jozsa *et al.*)

$$I \geq Q(\rho) - \sum_\alpha p_\alpha Q(\rho_\alpha)\,, \tag{9}$$

where the subentropy Q is defined as

$$Q = -\sum_{\alpha=1}^{M} \left(\prod_{\alpha \neq \beta} \frac{\omega_\alpha}{\omega_\alpha - \omega_\beta} \right) \omega_\alpha \ln \omega_\alpha\,. \tag{10}$$

If two or more of the eigenvalues are equal, Q remains finite if one takes the limit $\omega_\alpha \to \omega_\beta$.

• In order to control the weight of retained information during the RG procedure in a more rigorous way, the reduced density matrix of the left or right subsystem is written as the sum of the density matrices of mutually orthogonal mixed states belonging to the typical and atypical subspaces,

$$\rho^{(L)} = p_{\mathrm{typ}} \rho^{(L)}_{\mathrm{typ}} + (1 - p_{\mathrm{typ}}) \rho^{(L)}_{\mathrm{atyp}} \,, \tag{11}$$

where $\rho^{(L)}_{\mathrm{typ}}$ is formed from the M largest eigenvalues of $\rho^{(L)}$ and $\rho^{(L)}_{\mathrm{atyp}}$ from the remaining eigenvalues with $\mathrm{Tr}\rho^{(L)}_{\mathrm{typ}} = \mathrm{Tr}\rho^{(L)}_{\mathrm{atyp}} = 1$.

• The accessible information for such a binary channel would be less than the Kholevo bound

$$\chi \leq S(\rho) - p_{\text{typ}} S(\rho_{\text{typ}}) - (1 - p_{\text{typ}}) S(\rho_{\text{atyp}}) \,, \tag{12}$$

and larger than the Jozsa-Robb-Wootters lower bound

$$\chi \geq Q(\rho) - p_{\text{typ}} Q(\rho_{\text{typ}}) - (1 - p_{\text{typ}}) Q(\rho_{\text{atyp}}) \,. \tag{13}$$

• The schematic plot of the two bounds as a function of p_{typ}

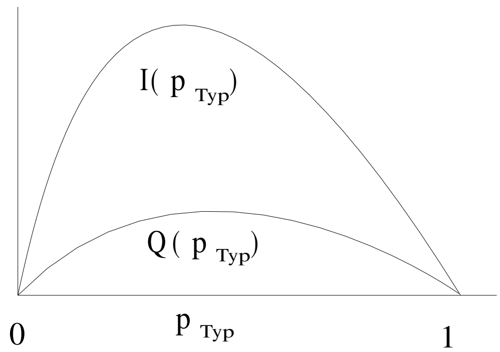


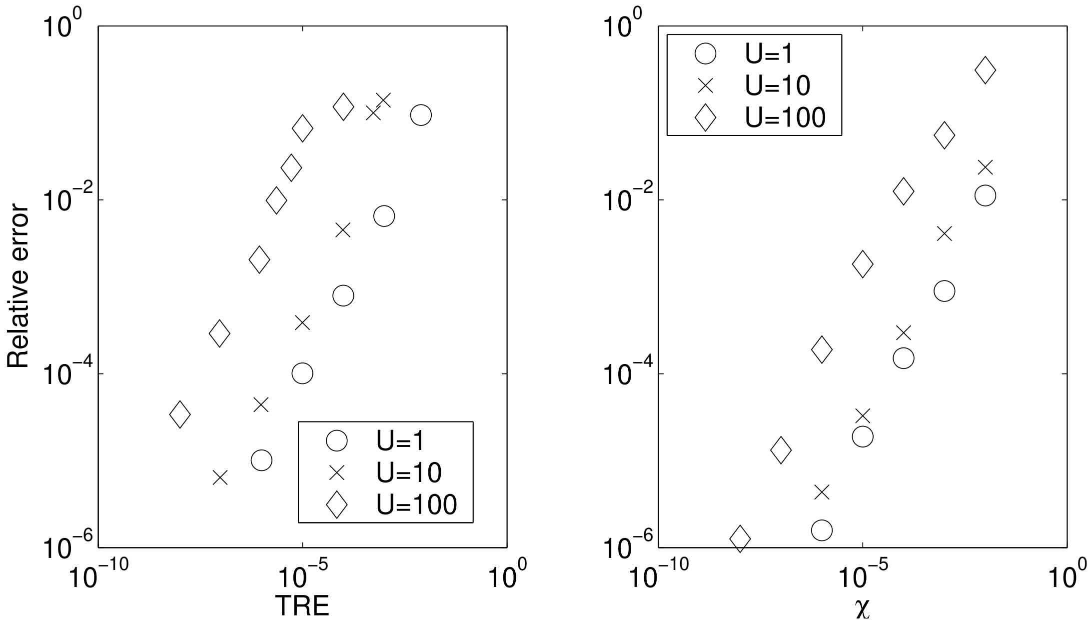
FIG. 1: Schematic plot of the upper and lower bonds on accessible information for a binary channel.

• In DMRG, however, the atypical subspace is neglected, which leads to a loss of information.

• For p_{typ} close to unity this loss is argued to be equal to the accessible information which is negligible if the probability of the atypical subspace is negligible.

• A modified DBSS procedure is proposed, where instead of fixing the truncation error, the Kholevo bound is required to be less than an ϵ fixed in advance, and the number of states M is chosen accordingly in every step.

• The relative error as a function of the truncation error and the threshold value of Kholevo's bound on accessible information obtained with the DBSS method and by Eq. (12), respectively.



The relative error is given as $(E_{\rm DMRG} - E_{\rm exact})/E_{\rm exact}$.

XI. RELATIONSHIP TO QUANTUM INFORMATION GENERATION

• For a renormalization step of the forward sweep

$$S_L(l) + S_{l+1} + I_L(l) = S_L(l+1)\,, \tag{14}$$

• the so-called mutual information $I_L(l)$ quantifies the correlation between the subsystem and the site.

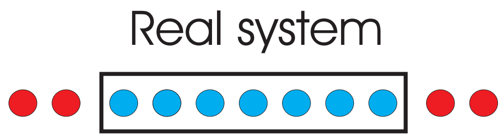
• The unitary operation applied on the basis states of the B_l • composite system is a type of LOCC it cannot increase the entanglement between B_L and B_R block.

• Therefore, the entropy of the B_l • composite system remains unchanged or decreased by forming the larger block B_L.

• The quantum information generated by the renormalization procedure of the forward sweep can be measured using Eq. (14) as

$$I_L(l) = S_L(l+1) - S_L(l) - S_{l+1}\,, \tag{15}$$

where l runs from 1 to $N-1$.



• If an effective system of length $N+2$ is formed by adding two non-interacting sites to the right end of the chain, all blocks containing 1 to N lattice sites of the original system can be formed by the forward sweep. The total information gain of a full half sweep can be calculated as $\Sigma_{l=1}^{N-1} I_L(l)$.

• If no truncation is applied, a sum rule holds, which relates the total information gain within a full half sweep and the sum of site entropies given as

$$\sum_{l=1}^{N-1} I_L(l) = -\sum_{l=1}^{N} S_l \, , \tag{16}$$

where we have used $S_L(1) = S_1$ and $S_L(N) = 0$.

• Due to the truncation of basis states, S_l changes for each half sweep as the method converges to the attractor renormalization process $S_L(l+1)$ is reduced to $S_L^{\mathrm{Trunc}}(l+1)$, thus $I_L(l)$ is also a function of subsequent sweeps.

• Once the DMRG method has converged, i.e., subsequent DMRG sweeps do not change $S_L(l)$ and S_l, the following equality should hold to a good accuracy

$$\sum_{l=1}^{N-1} I_L(l) \simeq -\sum_{l=1}^{N} S_l + \sum_{l=2}^{N} \left(S_L(l) - S_L^{\rm Trunc}(l) \right) . \tag{17}$$

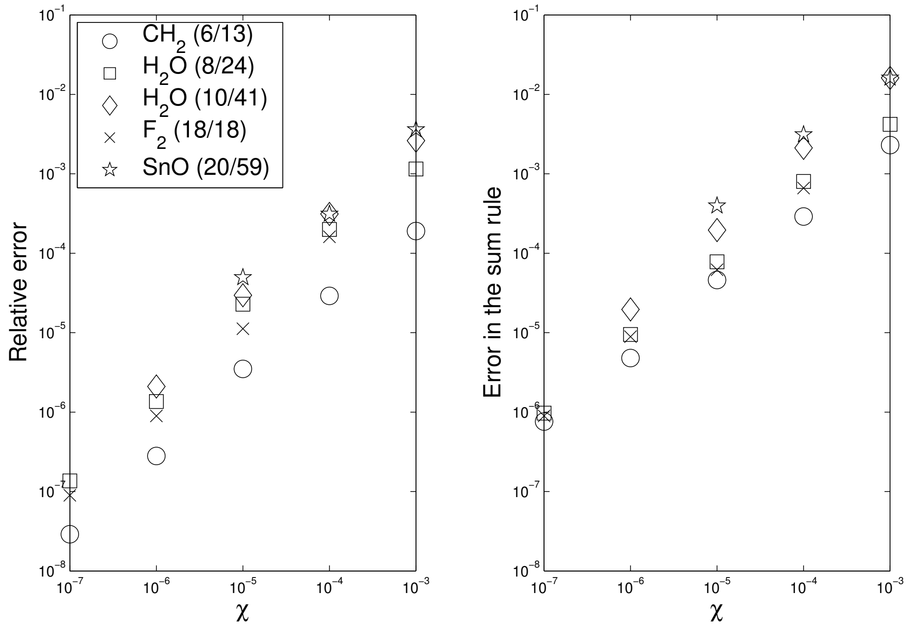
• We have found that when only the first few eigenvalues of ρ are retained, $p_{\rm typ}$ is already very close to unity.

• The loss of information could be equally controlled by requiring that $\chi \equiv S(\rho) - S(\rho_{\rm typ})$ should become less than ϵ in which case Eq. (17) leads to

$$\sum_{l=1}^{N-1} I(l) + \sum_{l=1}^{N} S_l < (N-1)\epsilon. \tag{18}$$

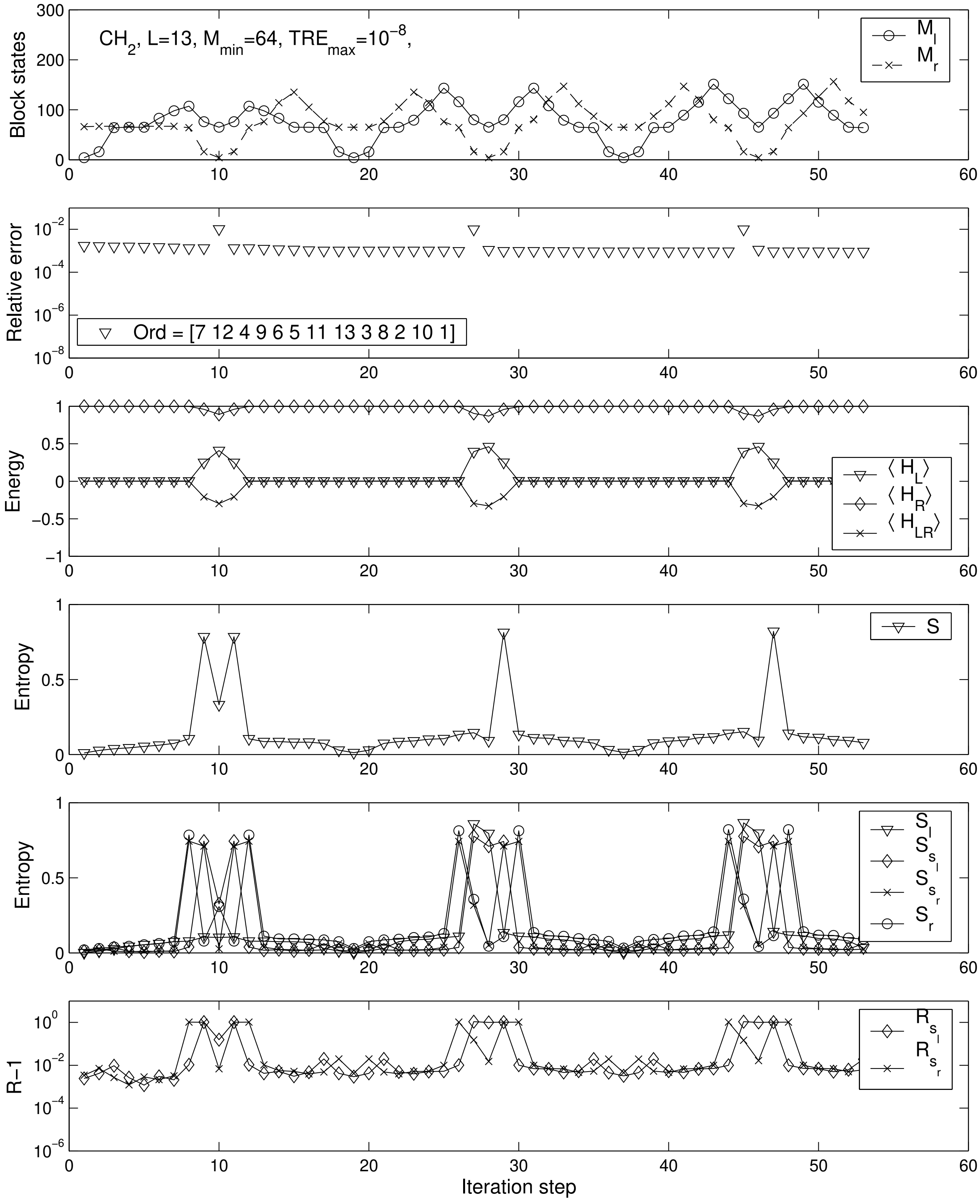
• Numerically we have found that the relative error and $\sum_l I(l) + \sum_l S_l$ are linear functions of χ on a logarithmic scale.

• The first panel shows the relative error as a function of χ, the second panel shows the error in the sum rule given by Eq. (18).



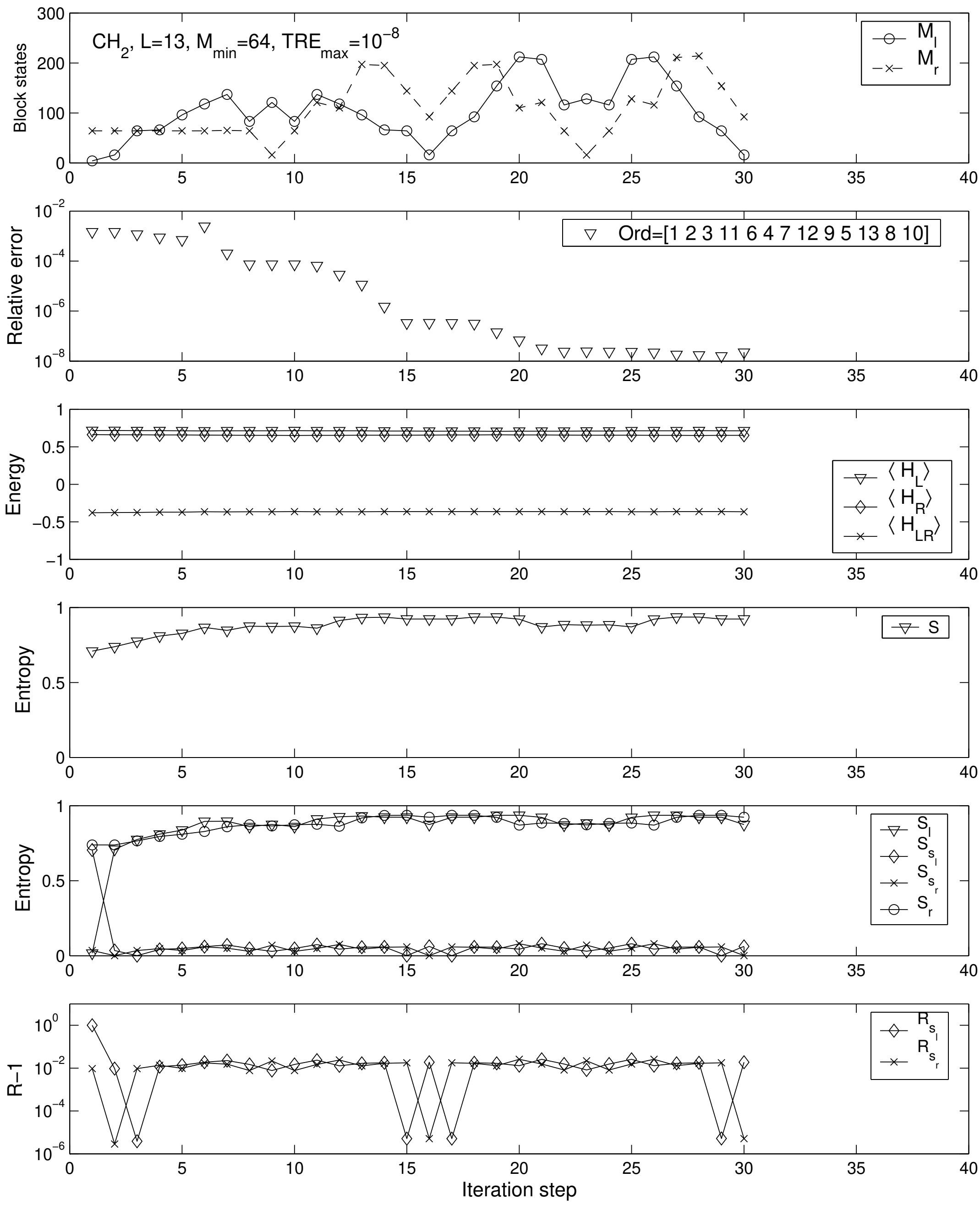
XII. ENTANGLEMENT LOCALIZATION VS. INTERACTION LOCALIZATION

- Analysis on a small molecule: CH_2 (6/13)
- It is hardly expected that DMRG would not converge.
- Ordering of orbitals with decreasing S_i values: $2\,,8\,,10\,,1\,,3\,,13\,,11\,,6\,,5\,,9\,,4\,,12\,,7$
- Example for an ordering when relevant information of the system is collected into one DMRG block giving rise to vanishing interblock interactions and very low entropy indicating high separability of the target state and lack of quantum information exchange between the blocks.

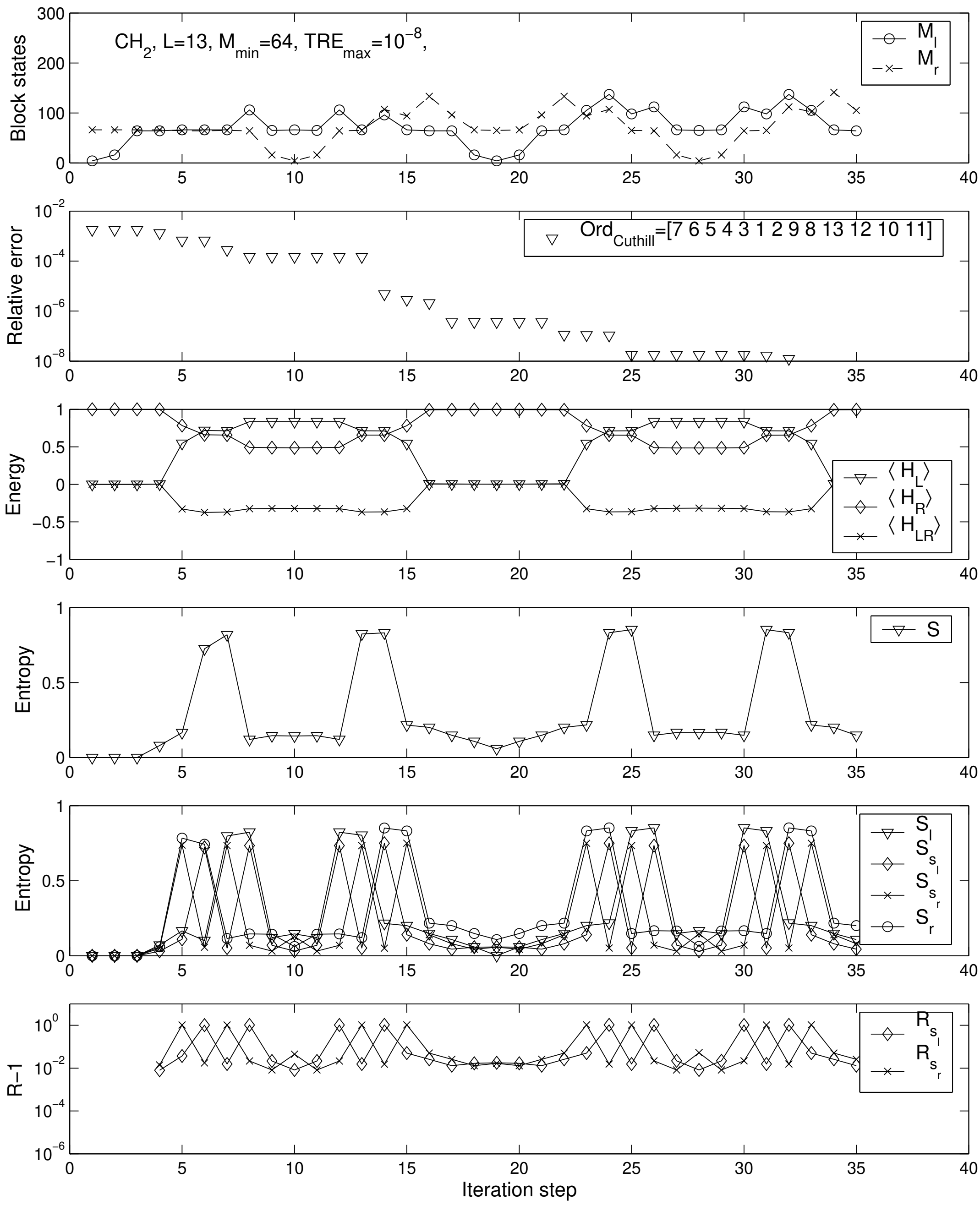


CH2, L=13, Mmin=64, TREmax=10−8,
Ml
Mr
Block states
Relative error
Ord = [7 12 4 9 6 5 11 13 3 8 2 10 1]
Energy
⟨HL⟩
⟨HR⟩
⟨HLR⟩
Entropy
S
Sl
Ssl
Ssr
Sr
R−1
Rsl
Rsr
Iteration step

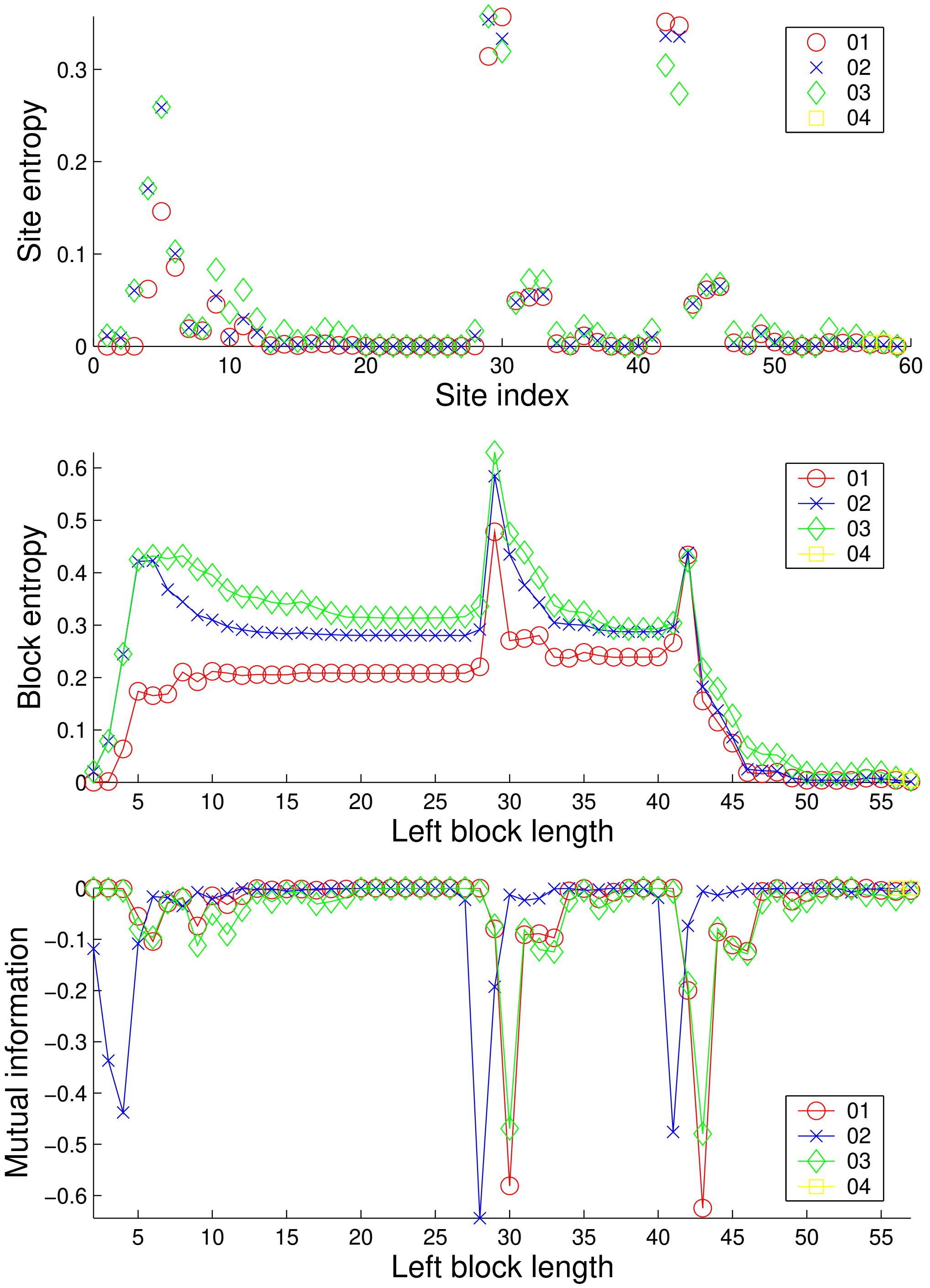
• The interaction between the blocks was maximized by putting the orbitals alternatively to the two ends with decreasing S_i.

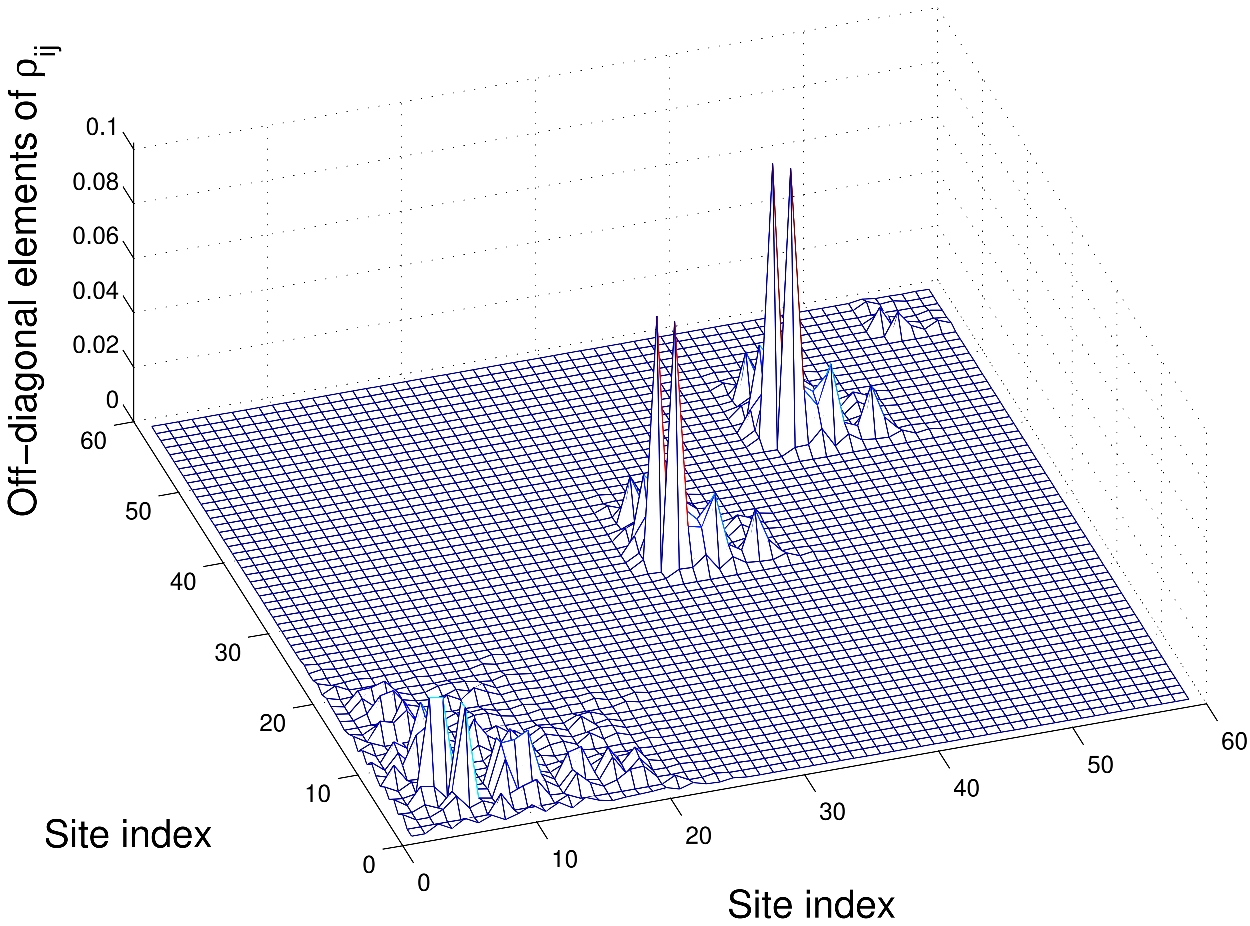


The block entropy profile was minimized by putting the orbitals alternatively to the center of the chain with decreasing S_i.



• Entanglement and interaction localization, Ex: SnO (20/59)




Off−diagonal elements of ρ
ij
0.1
0.08
0.06
0.04
0.02
0
60
50
40
30
20
10
0
Site index
0
10
20
30
40
50
60
Site index

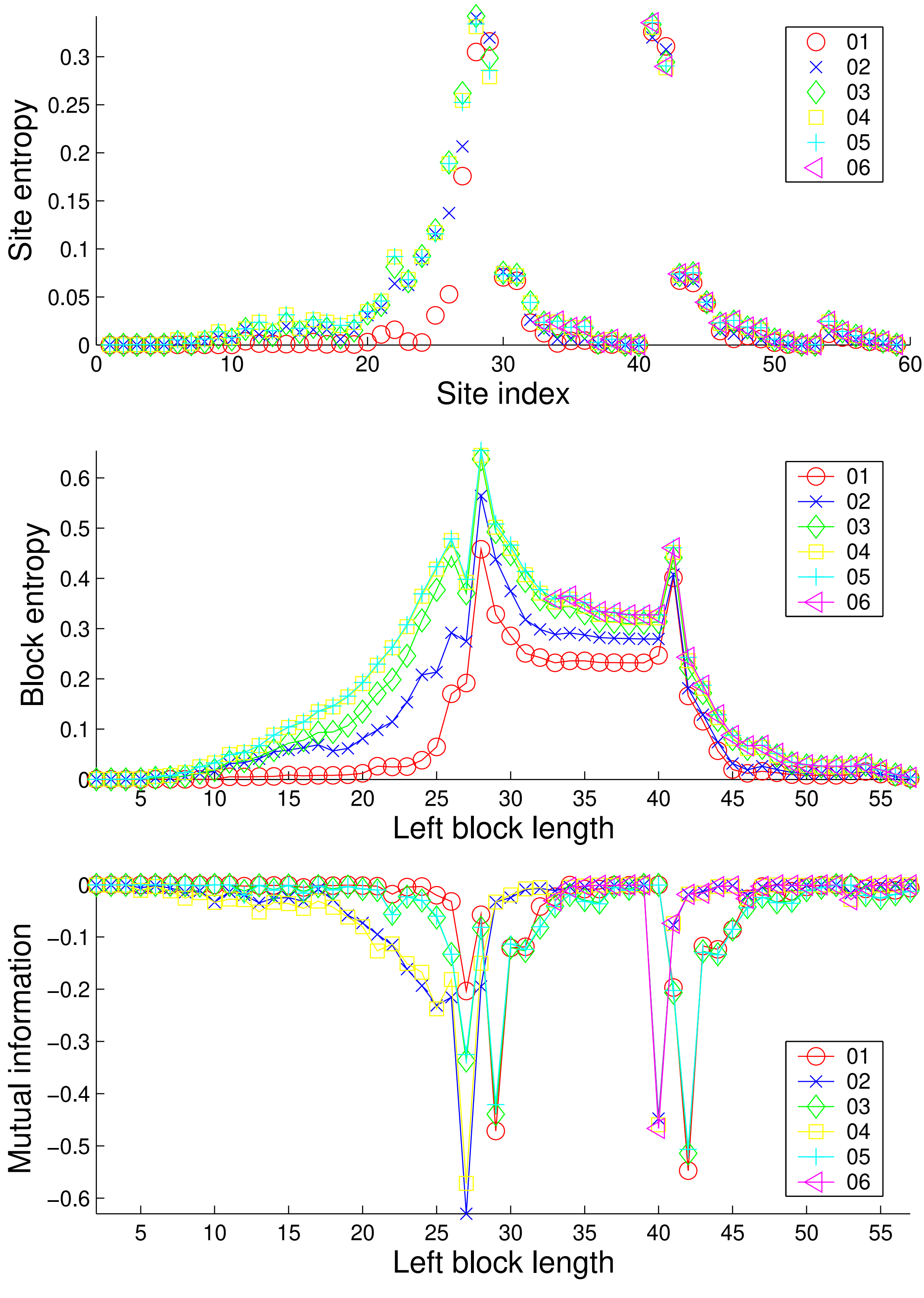
• Our proposed procedure to generate a good ordering:

Step 1. Using natural ordering of orbitals a quick DMRG calculation with some 100 block states is run and the site entropies and one-particle density matrix are determined.

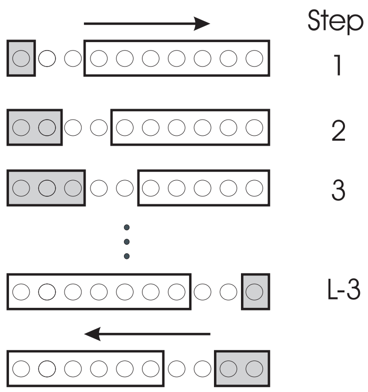
Step 2. If the one-particle density matrix is block diagonal, then orbitals are ordered in each symmetry sector with decreasing site entropies.

Step 3. Orbitals of each symmetry sector are placed next to each other with decreasing entropies counted from the center of the chain.

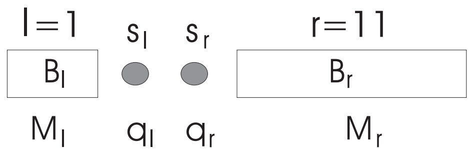
• This method guarantees a finite interblock coupling for several steps but also reduces the length of interaction and block entropy.



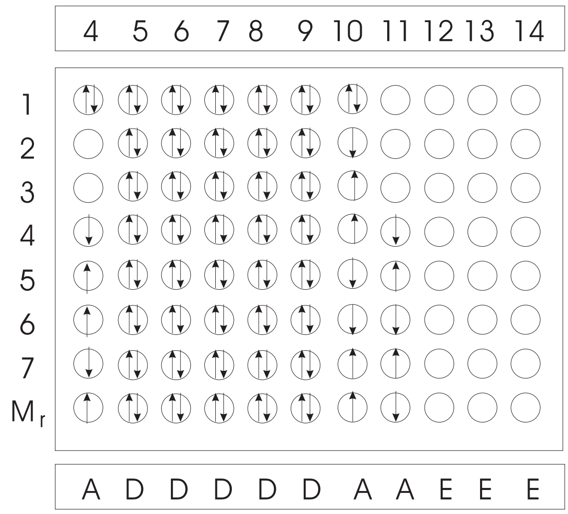
Site entropy
Site index
01
02
03
04
05
06
Block entropy
Left block length
01
02
03
04
05
06
Mutual information
Left block length
01
02
03
04
05
06



• The decomposition of the finite system to subsystems as a function of iteration steps corresponding to various partitionings. The shaded block represents the system block and the other block stands for the environment block. The arrows indicate the forward and backward sweeps.



$$l+1+1+r = 14$$



E = Empty sites
D = Doubly filled sites
A = Active sites

- After working out algebra one gets:

$A_L^{(4)}$

$A_L^{(1)}$

$A_R^{(4)} \rightarrow A_R^{(4)\prime}$

$A_L^{(3)} \rightarrow A_L^{(3)\prime}$

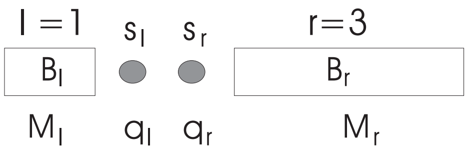
$A_R^{(1)} \rightarrow A_R^{(1)\prime}$ only on AS

$A_R^{(2)} \rightarrow A_R^{(2)\prime}$ only on AS

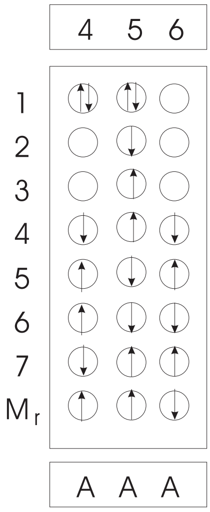
$A_R^{(3)} \rightarrow A_R^{(3)\prime}$

$A_L^{(2)} \rightarrow A_L^{(2)\prime}$ and $A_L^{(4)\prime\prime}$

- AS vector: ordering of sites with decreasing S_i.

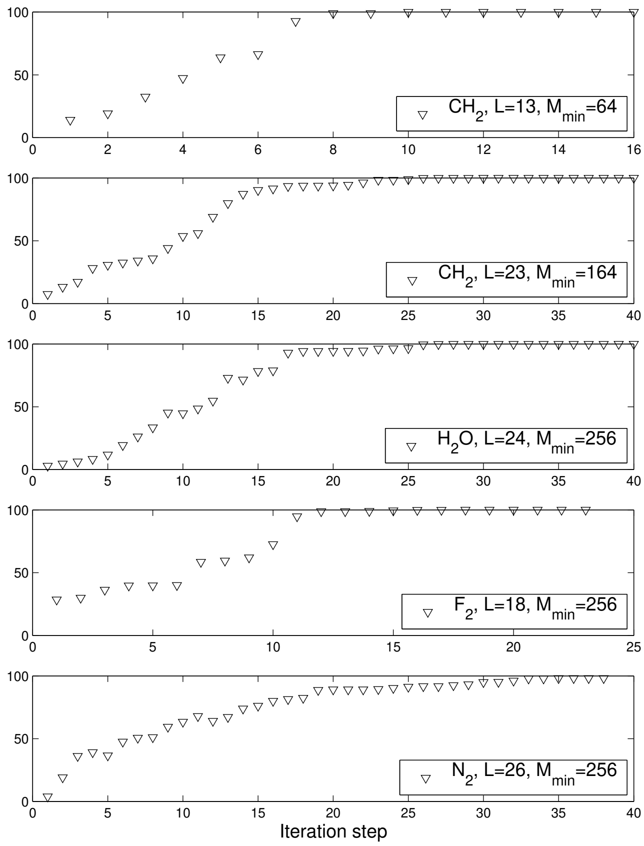


$$l+1+1+r = 6$$



A = Active sites

Percent of correlation energy as a function of iteration step obtained by DEAS procedure for the CH_2, H_2O, N_2, F_2 molecules.



XIV. ? FUTURE POSSIBILITIES ?

- Relativistic 4-component quantum chemistry DMRG (u.p.)
- Dynamically reordered orbitals (DRO) procedure (u.p.)
- DMRG study of QIT using non-orthogonal basis states
- Entanglement distillation and purification procedures
- Quantum data compression (QDC)
- Quantum error correction codes (QECC) simulation
- Multilevel Quantum Cryptography, Quantum key distribution
- Quantum computer simulation
- Genetics, coding, decoding