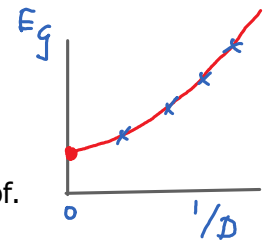
Tangent space methods III: Energy variance, controlled bond expansion (CBE)

1. Energy variance [Hubig2018]



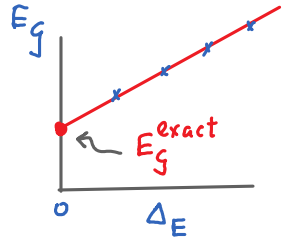
When doing MPS computations involving SVD truncations of virtual bonds, the results should be computed for several values of the bond dimension, D, to check convergence as $D \to \infty$. Often it is also necessary to extrapolate the results to $D = \infty$, e.g. by plotting results versus $1/D$ or some power thereof.

However, for some computational schemes, it is not *a priori* clear how the observable of interest scales with D, nor how it should be extrapolated to $D = \infty$. An example is ground state energy when computed using 1-site DMRG with subspace expansion [Hubig2015], because it does not rely on SVD truncation of bonds.

Thus, it is of interest to have a reliable error measure without requiring costly 2-site DMRG. A convenient scheme was proposed in [Hubig2018], based on a smart way to approximate the full energy variance,

$$\Delta_E := \|(H - E)\psi\|^2 = \langle\psi|(\hat{H} - E)^2|\psi\rangle \quad \text{(= zero for an exact eigenstate)} \quad (1)$$

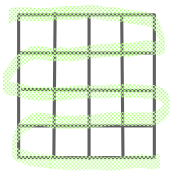
$$= \langle\psi|\hat{H}^2|\psi\rangle - E^2, \quad \text{with} \quad E = \langle\psi|\hat{H}|\psi\rangle \quad (2)$$



Then extrapolations can be done by computing quantity of interested for several D, but plotting the results via Δ_E, and extrapolating to $\Delta_E \to 0$

If quantity of interest is energy, then extrapolation is linear, $E_g(\Delta_E) = E_g^{exact} + a \cdot \Delta_E$ (3)

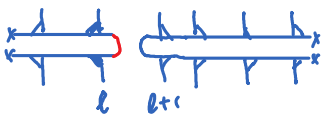
Computing $\langle\psi|\hat{H}^2|\psi\rangle$ directly is costly for large systems with long-ranged interactions, such as 2D systems treated by DMRG snakes. Also, computing Δ_E as the difference between two potentially large numbers is prone to inaccuracies. [Hubig2018] found a computation scheme in which the subtraction of such large numbers is avoided *a priori*.

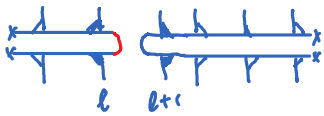


Key idea: use projectors $P^{n\perp}$ onto mutually orthogonal, irreducible spaces $\mathbb{V}^{n\perp}$

Recall (2.11): $\mathbb{1}_{\mathbb{V}} = \mathbb{1}_d^{\otimes \mathcal{L}} = \sum_{n=0}^{\mathcal{L}} P^{n\perp}$, completeness (4) $\qquad P^{n\perp} P^{n'\perp} = \delta^{nn'} P^{n\perp}$ orthogonality (5)

with $P^{0\perp} = |\Psi\rangle\langle\Psi|$ (6)

(2.16) $P^{1\perp} = \sum_{\ell=1}^{\mathcal{L}}$ 

, (2.17) $P^{2\perp} = \sum_{\ell=1}^{\mathcal{L}}$ 

(7)

Insert completeness into definition of variance: $\Delta_E \overset{(4)}{=} \langle\psi|(\hat{H} - E)\sum_{n=0}^{\mathcal{L}} P^{n\perp}(\hat{H} - E)|\psi\rangle =: \sum_{n=0}^{\mathcal{L}} \Delta_E^{n\perp}$ (8)

Now two crucial simplifications occur:

Now two crucial simplifications occur:

$$\Delta_E^{0\perp} \overset{(5)}{=} \langle\psi|(\hat H - E)\underbrace{|\psi\rangle\langle\psi|}_{(6)\ P^{0\perp}}(\hat H - E))|\psi\rangle = (E-E)(E-E) = 0 \tag{9}$$

largest contribution to variance cancels by construction!

$n>0$

$$\Delta_E^{n\perp} = \langle\psi|(\hat H - E)P^{n\perp}(\hat H - E))|\psi\rangle = \langle\psi|\hat H \underbrace{P^{n\perp}}_{= P^{n\perp}P^{n\perp}}\hat H|\psi\rangle, \quad \text{since } P^{(n>0)\perp}|\psi\rangle \overset{(5,6)}{=} 0 \tag{10}$$

(TS-II.2.11)

$$= \| P^{n\perp}\hat H\psi\|^2 \tag{11}$$

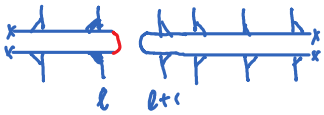
In practice, approximate Δ_E by the first two nonzero terms:

$$\Delta_E \simeq \Delta_E^{2s} = \Delta_E^{1\perp} + \Delta_E^{2\perp} = \boxed{\langle\psi|\hat H P^{1s}\hat H|\psi\rangle + \langle\psi|\hat H P^{2s}\hat H|\psi\rangle} \tag{12}$$

(11) is exact if longest-range terms in $\hat H$ are nearest-neighbor, because then $P^{(n\geq 3)\perp}\hat H|\Psi\rangle = 0$ (13)

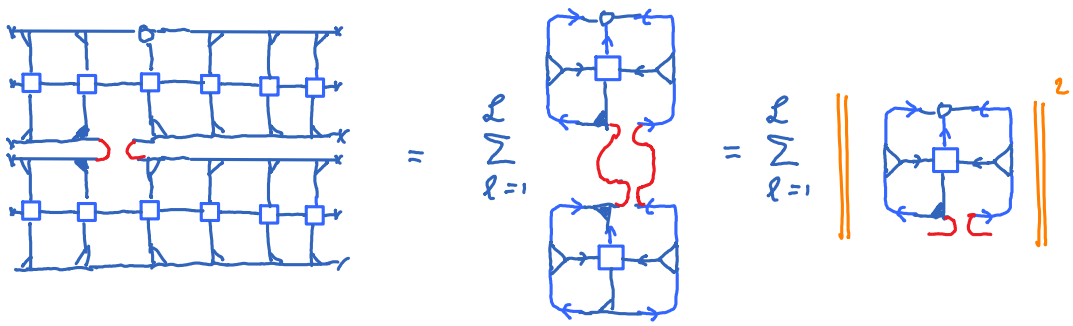
[Gleis2022a]

Explicit computations:

$n=1$: Recall (TS-II.2.16) $P^{1\perp} = \sum_{\ell=1}^{\mathcal{L}}$  $= \sum_{\ell=1}^{\mathcal{L}} P^{DK}_{\ell,\ell+1}$ (14)

mutually orthogonal! (TS-I.4.15)

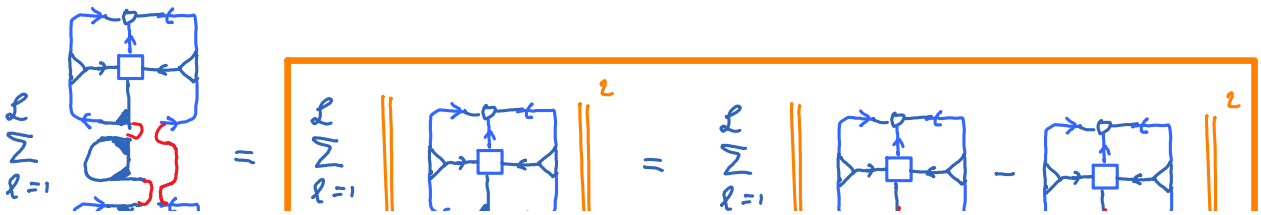
$$\Delta_E^{1\perp} \overset{(10)}{=} \langle\psi|\hat H \underbrace{P^{1\perp}P^{1\perp}}\hat H|\psi\rangle \overset{(14)}{=} \sum_{\ell\ell'=1}^{\mathcal{L}} \langle\psi|\hat H \overbrace{P^{DK}_{\ell,\ell+1}P^{DK}_{\ell',\ell'+1}}^{\sim\delta_{\ell\ell'}}\hat H|\psi\rangle = \sum_{\ell=1}^{\mathcal{L}} \| P^{DK}_{\ell,\ell+1}\hat H\psi\|^2 \tag{15}$$

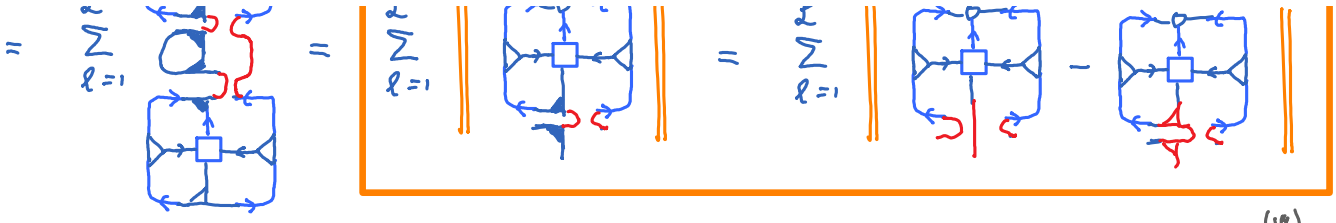
$=$  (16)

We would like to avoid computing  explicitly, because of its large image dimension.

So rewrite, using isometry condition for discarded sector:  (17)

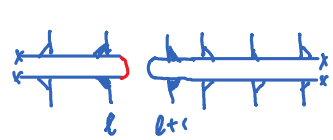
and completeness of kept together with discarded isometries:  (18)

$=$ 



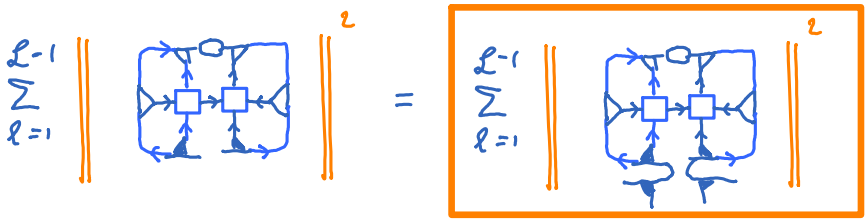
(19)

$n = 2$: Recall (2.17)



(20)

$$\Delta_E^{2\perp} = \langle\psi|\hat{H}\,\overbrace{P^{2\perp}}^{P^{2\perp}P^{2\perp}}\,\hat{H}|\psi\rangle = \|P^{2\perp}H\psi\|^2 = \sum_{\ell=1}^{\mathcal{L}-1}\|P^{D}_{\ell,\ell+1}H\psi\|^2 \qquad (21)$$



(22)

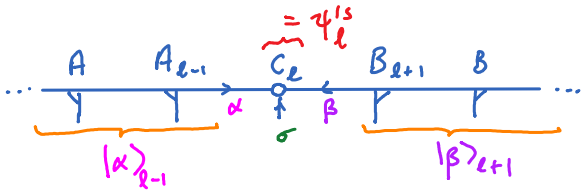
again use

(23)

2. Controlled bond expansion (CBE) for DMRG [Gleis2022]

Problem: when exploiting symmetries, 1-site DMRG performs poorly, because it does not explore subspaces with different quantum numbers. An early remedy for this is 2-site DMRG, but that is computationally much more expensive than 1-site DMRG. Subsequent suggestions for 1-site DMRG with symmetries are 'density matrix perturbation' [White2005], the 'center matrix wave function formalism [McCulloch2007], 'subspace expansion' [Hubig2015], and 'controlled bond expansion' (CBE) [Gleis2022], which performs best.

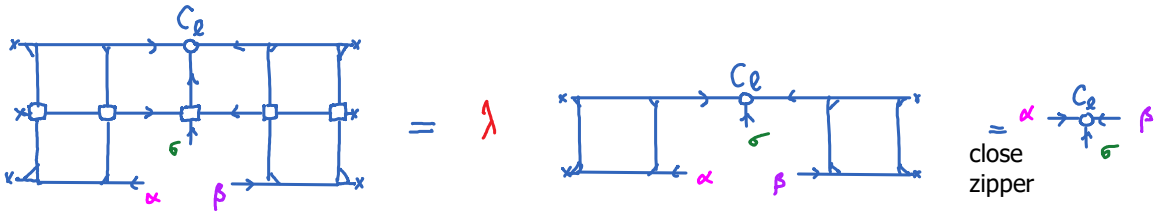
Reminder of 1-site DMRG,
in site-canonical representation:
Local basis: $|\alpha,\sigma,\beta\rangle := |\alpha\rangle_{\ell-1}|\sigma\rangle|\beta\rangle_{\ell+1}$



(1)

Minimize energy with constraint of fixed normalization, 1 site at a time:

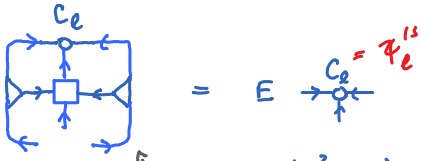
$$\frac{\partial}{\partial C_\ell^\dagger}\left[\langle\psi|\hat{H}|\psi\rangle - \lambda\langle\psi|\psi\rangle\right] = 0 \qquad (2)$$



(3)

$$[H_\ell^{1s}]^{a'}{}_{a}[\psi_\ell^{1s}]^{a} = E[\psi_\ell^{1s}]^{a}$$

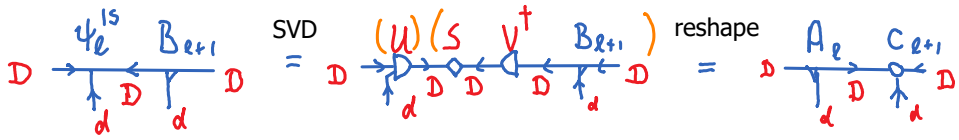
$a = (\alpha,\sigma,\beta)$



(4)

cost ~ $\mathcal{O}(D^3 d w)$

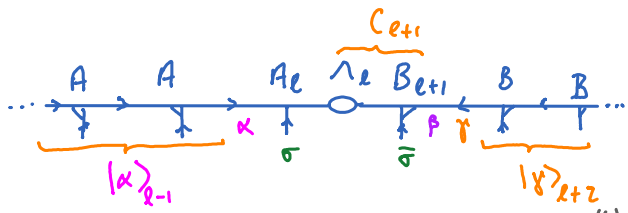
Solve for 'eigenvector' with lowest eigenvalue, say $\tilde{\psi}_\ell^{1s}$, then do SVD on it to move to next site:



(5)

Important: dimensions of $\tilde{C}$ are <u>fixed</u>, hence truncation is neither needed nor possible!

Reminder of 2-site DMRG,
in site-canonical representation:
Local basis: $|\alpha,\sigma,\bar{\sigma},\gamma\rangle := |\gamma\rangle_{\ell+2}|\bar{\sigma}\rangle|\sigma\rangle|\alpha\rangle_{\ell-1}$

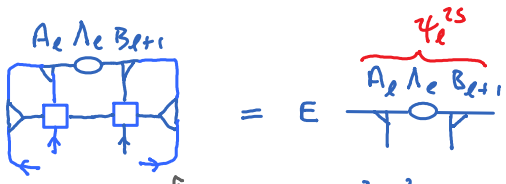


(6)

Minimize energy two sites at a time:

$$[H_\ell^{2s}]^{a'}{}_{a}[\psi_\ell^{2s}]^{a} = E[\psi_\ell^{2s}]^{a}$$

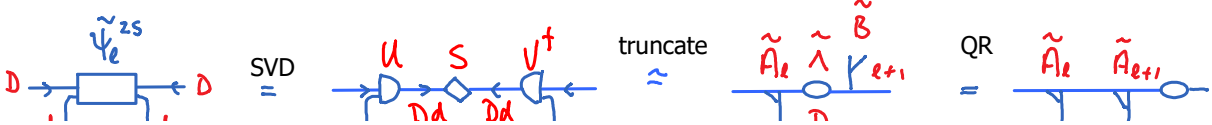
$a = (\alpha,\sigma,\bar{\sigma},\gamma)$

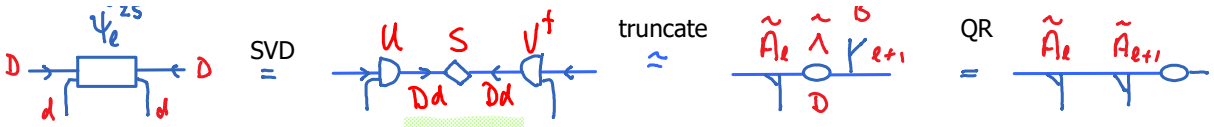


(7)

cost ~ $\mathcal{O}(D^3 d^3 + D^3 d^2 w)$

Solve for 'eigenvector' with lowest eigenvalue, $(\widetilde{MB})$, then do SVD and <u>truncate</u> (!) to move to next site:





(8)

Problem of single-site optimization: it is constrained to a variational space defined by outgoing state spaces $|\alpha\rangle_{\ell-1}|\beta\rangle_{\ell+1}$. If the ranges of quantum numbers Q_α and Q_β for these spaces are too small to accurately represent the ground state, single-site DMRG has no way to enlarge them.

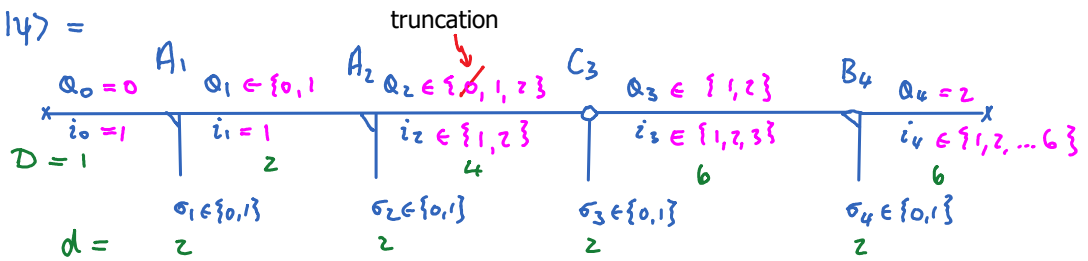
Two-site optimization does not have this problem: the <u>action of H</u> on two sites <u>enlarges</u> bond dimension in between, adding the full range of quantum numbers needed on that bond. If a certain quantum number was missing on that bond before the action of H, but appears afterwards with non-negligible weight, it will survive after SVD and truncation. Hence: two-site optimization can <u>add</u> missing quantum numbers, if needed.

But this comes at a cost: effective two-site Hamiltonian has dimension $D^2d^2 \times D^2d^2$.
By contrast, effective one-site Hamiltonian has dimension $D^2d \times D^2d$. (9)

<u>Example:</u> 4-site chain of spinless fermions, with total charge $Q_4 = 2$

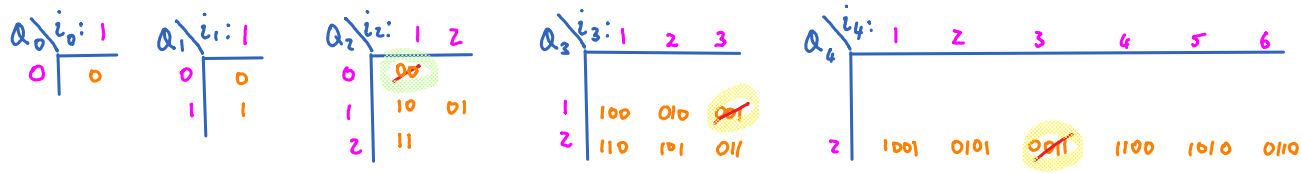
exploiting charge quantum numbers, with conservation law $Q_\ell = Q_{\ell-1} + \sigma_\ell$, where $\sigma_\ell \in \{0, 1\}$ (empty site, occupied site)

i_ℓ enumerates distinct states with same charge (10)



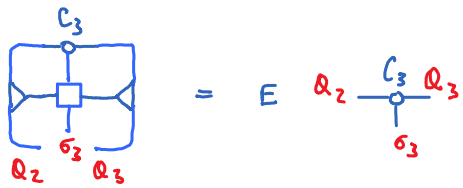
(11)

states spaces: (12)



truncation on bond $\ell = 2$ causes missing states on all later bonds

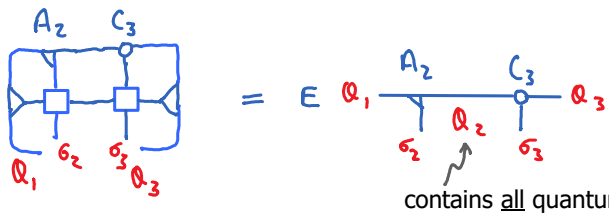
1s DMRG:



updated C_3 has <u>same</u> bond dimensions as initial C_3 (13)

1-site optimization of truncated C_3 will never find a good ground state if latter has non-negligible contributions from missing blocks.

2s DMRG:



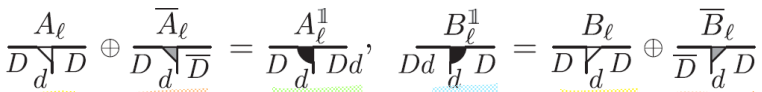
updated A_2, C_3 can have <u>larger</u> bond dimensions than initial A_2, C_3 (14)

contains <u>all</u> quantum numbers consistent with $Q_1 + \sigma_2 = Q_2 = Q_3 - \sigma_2$

2-site optimization can reinstate missing blocks!

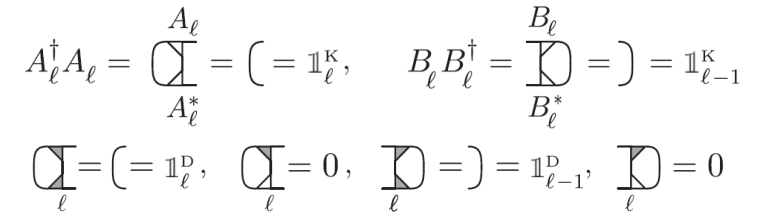
Which part of 2-site space is missed by H^{1s}_{ℓ} and $H^{1s}_{\ell+1}$?

Recall kept+discarded decomposition: $D + \bar{D} = dD$, $\bar{D} = D(d-1)$



(15)

Orthonormality of kept and discarded isometries:



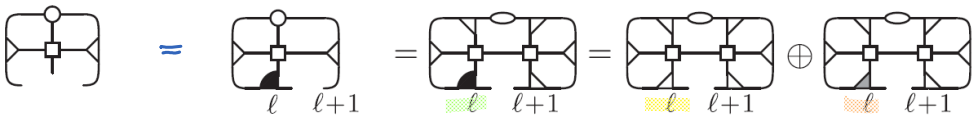
(16a)

(16b)

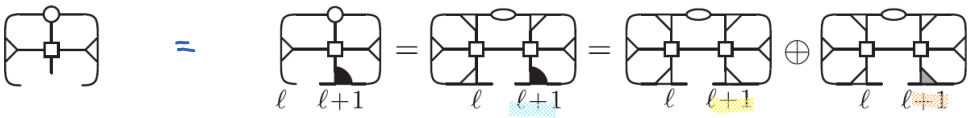
Completeness:

(16c)

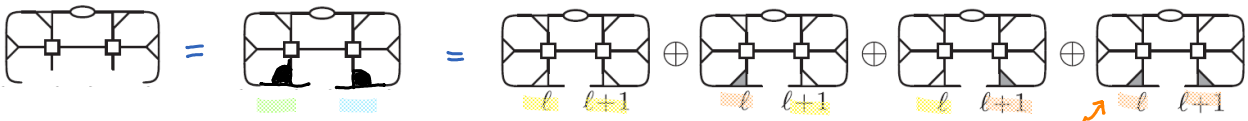
Compare action of 1-site and 2-site Hamiltonians:



(17a)



(17b)



(17c)

This part is not captured by 1s DMRG

This can also seen by considering energy variance:

$$\Delta_E^{1\perp} = \sum_{\ell=1}^{\mathscr{L}} \left\| \ell \right\|^2 = \| P^{1\perp} H \psi \|^2 , \qquad (18a)$$

Minimized by 1s DMRG, vanishes for converged 1s-GS.

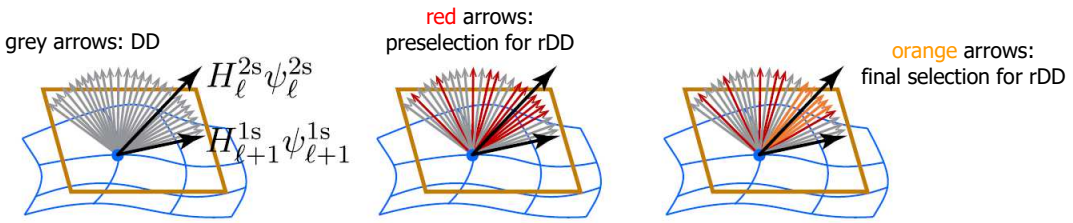
$$\Delta_E^{2\perp} = \sum_{\ell=1}^{\mathscr{L}-1} \left\| \ell \quad \ell+1 \right\|^2 = \| P^{2\perp} H \psi \|^2 \qquad (18b)$$

Minimized by 2s DMRG, vanishes for converged 2s-GS.

Subspace missed by 1s DMRG but explored by 2s DMRG is the DD subspace: image $\left(\overline{A}_\ell \otimes \overline{B}_{\ell+1} (\mathbb{1} \otimes \mathbb{1})\right)$

It contains 'missing' symmetry sectors (good!), but is huge (bad!) dimension: $\bar{D} \times \bar{D}$ = huge! (19)

Key insight [Gleis2022]: $H^{2s}_{\ell} \psi^{2s}_{\ell}$ has significant weight only on <u>small</u> subspace of DD, the 'relevant DD' (rDD) !



(20)

Hence, it suffices to expand bond to include only the rDD ! View rDD as image $\left(\widetilde{A}^{\mathrm{tr}}_{\ell} \otimes \overline{B}_{\ell+1} (\mathbb{1} \otimes \mathbb{1})\right)$

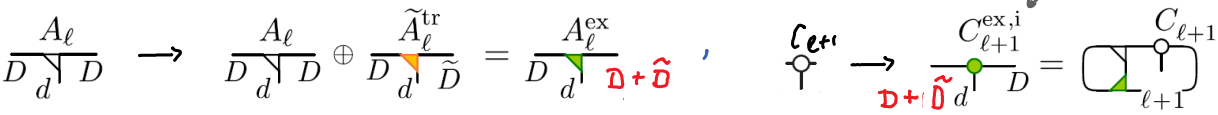
or image $\left(\overline{A}_{\ell} \otimes \widetilde{B}^{\mathrm{tr}}_{\ell+1} (\mathbb{1} \otimes \mathbb{1})\right)$

Truncated isometries $\mathbb{1}$ or $\mathbb{1}$ can be found via 'shrewd selection' = (i) preselection, then (ii) final selection (see next section)

Controlled bond expansion (CBE) [for right-to-left sweep]

(i) Compute truncated isometry $\widetilde{A}^{\mathrm{tr}}_{\ell}$ ( )

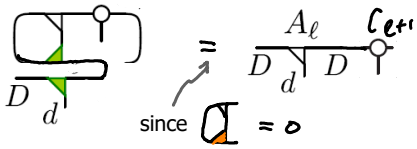
(ii) expand bond ℓ :

replace 

(21)

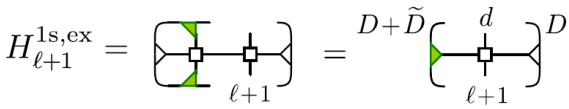
so that initialized version of expanded bond = old bond:

$$A_\ell^{exp} C_\ell^{exp} = A_\ell C_\ell$$



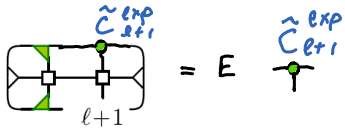
(22)

and construct expanded 1s Hamiltonian:

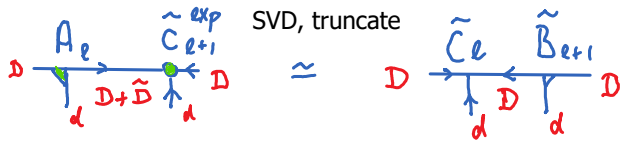

(23)

(iii) Find GS of expanded 1s Hamiltonian:

(e.g. Lanczos eigensolver), as in 1s DMRG:


(24)

(iv) Shift isometry center from $\ell+1$ to ℓ :



(25)

The truncated weight at step (iv), say ξ , serves as error measure.

In practice: suppose we want to gradually grow the bond dimension by a factor α per sweep. Then, for each update, we need to increase bond dimension from an initial D_i to a final $D_f = \alpha D_i$, with $\alpha > 1$. (26)

Thus, expand from D_i to $D_i + \widetilde{D} = D_f(1+\delta)$, with $\delta > 0$ (27)

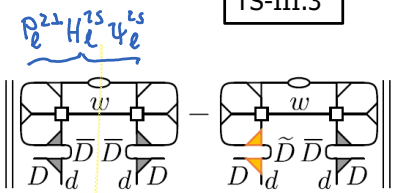
and in (iv), truncate from $D_f(1+\delta)$ to D_f . (28)

Typical choices: $\alpha = 1.1$, $\delta = 0.1$. (29)

3. Shrewd selection

Goal: truncate $\overline{A}_\ell(\overset{\overline{D}}{\blacktriangledown}) \to \widetilde{A}^{\mathrm{tr}}_\ell(\overset{\widetilde{D}}{\blacktriangledown})$ to minimize $C_1 =$

orthogonal complement — truncated complement



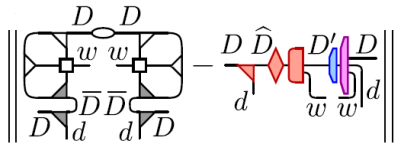
(1)

Optimal truncation can be achieved via SVD; but that has 2s costs, $\mathcal{O}(D^3 d^3)$

Instead, use `shrewd selection' (cheap, efficient, practical, though not strictly optimal), involving two steps:

(i) Preselection: truncate $\overline{A}_\ell(\overset{\overline{D}}{\blacktriangledown}) \to \widehat{A}^{\mathrm{pr}}_\ell(\overset{\widehat{D}=D}{\blacktriangledown})$ to minimize $C_2 =$

orthogonal complement — preselected complement

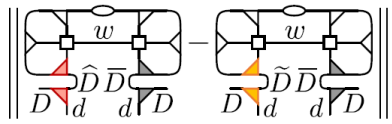


(2)

Truncate central bond in presence of its environment, with MPO bond open (to reduce numerical costs)

(ii) Final selection: truncate $\widehat{A}^{\mathrm{pr}}_\ell(\overset{\widehat{D}}{\blacktriangledown}) \to \widetilde{A}^{\mathrm{tr}}_\ell(\overset{\widetilde{D}<D}{\blacktriangledown})$ to minimize $C_3 =$

preselected complement — truncated complement

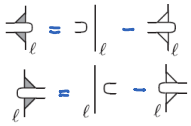


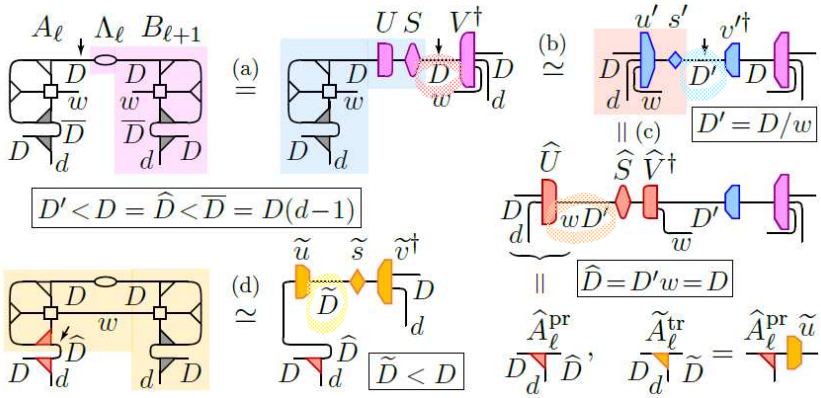
(3)

Truncate again, now in with MPO bond closed, as appropriate for $H_\ell^{2s} \psi_\ell^{2s}$

Details of preselection [steps (a-c)] and final selection [step (d)]:

- arrows indicate bond being opened before doing SVD
- shading and symbols in matching colors indicate SVD input and output
- output is written as $USV^\dagger$ or $usv^\dagger$ when involving no or some truncation, respectively
- use





(a) Canonicalize right side (shaded pink) of diagram, assigning its weights to central MPS bond.

(b) Truncate central MPS bond, $D \to D' = D/w$ (reason for this choice: see (d))

(c) Regroup, to combine truncated MPS bond and MPO bond into composite bond of dimension $\widehat{D} = D'w = D$

If using exact arithmetic, this would involve no truncation. In practice (numerically) zero singular values $\mathcal{O}(10^{-16})$, may arise. They must be truncated to ensure $= 0$, so that image $(\blacktriangledown) \subset$ image $(\blacktriangledown)$.

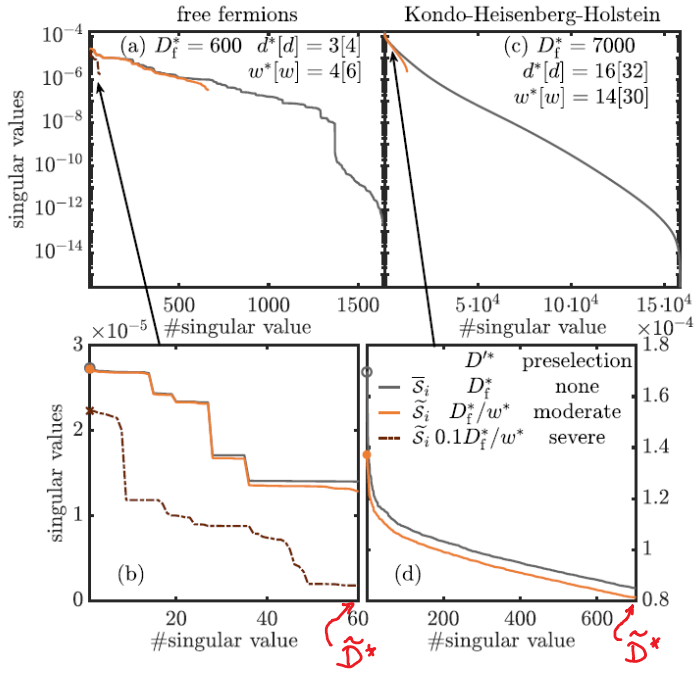
(d) Final selection: close MPO bond, then truncate central MPS bond: $\widehat{D} \to \widetilde{D} < D$ (e.g. $\widetilde{D} = 0.1 D$).

To ensure 1s costs for this step, we need $D'w = \widehat{D} = D$, hence choose $D' = D/w$ in (b).

Important: By design, every step has at most 1s costs, $\mathcal{O}(D^3 d w)$

Moreover, CBE captures the most most relevant contributions from H_ℓ^{2s}

$\Rightarrow$ 2s accuracy and convergence per sweep, at 1s cost !!



Comparison of three truncation settings:

grey: optimal truncation via SVD (grey)

serves as a reference

orange: moderate preselection, $D'^* = D_f^*/w^*$

then final selection $\tilde{D}^* = 0.1\, D_f$

agrees rather well with reference!

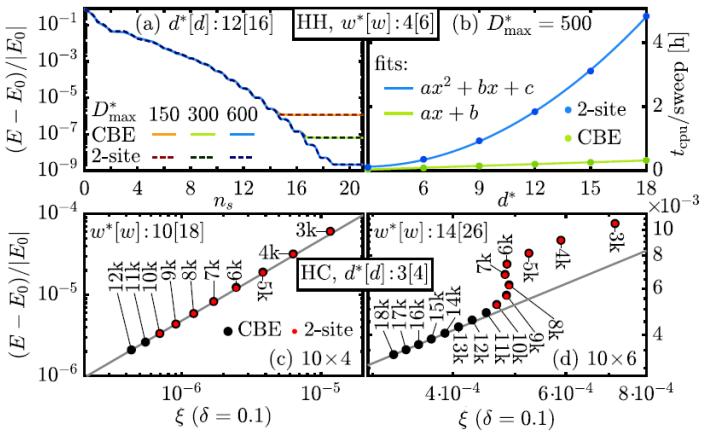
brown: severe preselection, $D'^* = 0.1\, D_f^*/w^*$

then final selection $\tilde{D}^* = 0.1\, D_f$

misses some information from reference

Take-home message: optimal truncation requires computation of a huge amount of singular values, most of which are discarded anyway. Those that are kept can be very well captured using shrewd selection!

Results for CBE-DMRG:



(a) CBE and 2s DMRG have same convergence rate per sweep.

(b) CBE has 1s costs ~ $\mathcal{O}(d)$, much faster than 2s DMRG ~ $\mathcal{O}(d^2)$

(c,d) Reliable convergence with increasing D^*, decreasing ξ

FIG. 3. Hubbard-Holstein (HH) model: (a) Convergence of the GS energy versus number of half-sweeps n_s at fixed $d^* = 3(N_{ph}^{max} + 1)$. E_0 was obtained by linear ξ extrapolation of data from $D^*_{max} \in [1000, 1200]$. (b) CPU time per sweep for various d^* at fixed D^*_{max}, showing d^* (CBE) vs d^{*2} (2s) scaling. Hubbard cylinders (HC): Error in GS energy vs ξ for (c) 10×4 and (d) 10×6 HCs, obtained with CBE (black) and 2s (red) DMRG, for various D^*_{max} (legends). Since 2s CPU times far exceed those of CBE, 2s data is only shown for $D^*_{max} \leq 10k$. Reference energies $E_0 = -27.881\,694\,2$ (10×4) and $-41.747\,496\,1$ (10×6) are obtained by linear ξ extrapolation of the four most accurate CBE results to $\xi = 0$ (gray line).

$$H_{HH} = -\sum_{\ell\sigma}(c^\dagger_{\ell\sigma}c_{\ell+1\sigma} + \text{H.c.}) + 0.8\sum_{\ell} n_{\ell\uparrow}n_{\ell\downarrow}$$
$$+ 0.5\sum_{\ell} b^\dagger_\ell b_\ell + \sqrt{0.2}\sum_{\ell}(n_{\ell\uparrow} + n_{\ell\downarrow} - 1) \times (b^\dagger_\ell + b_\ell)$$

$$H_{HC} = -\sum_{\langle \ell,\ell'\rangle,\sigma}(c^\dagger_{\ell\sigma}c_{\ell'\sigma} + \text{H.c.}) + 8\sum_{\ell} n_{\ell\uparrow}n_{\ell\downarrow}$$

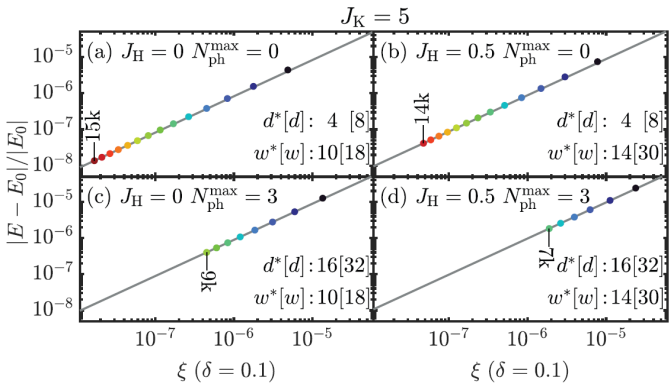


FIG. S-10. Error in GS energy versus discarded weight for the Kondo-Heisenberg-Holstein (KHH) model on a 10×4 cylinder, with (a) only Kondo coupling, (b) Kondo and Heisenberg coupling, (c) Kondo and Holstein coupling and (c) Kondo, Heisenberg and Holstein coupling. Legends state our choices for J_H and N_{ph}^{max}, and corresponding values of $d^*[d]$ and $w^*[w]$.

$$H_{KH} = -\sum_{\langle \ell,\ell'\rangle,\sigma}(c^\dagger_{\ell\sigma}c_{\ell'\sigma} + \text{H.c.}) + J_K\sum_{\ell} \mathbf{S}_\ell \cdot \mathbf{s}_\ell + \frac{1}{2}\sum_{\langle \ell,\ell'\rangle} \mathbf{S}_\ell \cdot \mathbf{S}_{\ell'}$$

J_K tunes quantum phase transition between two phases with different Fermi surface volumes.

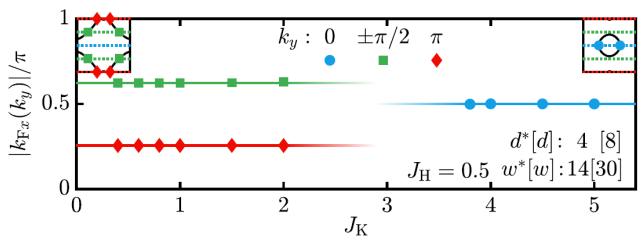
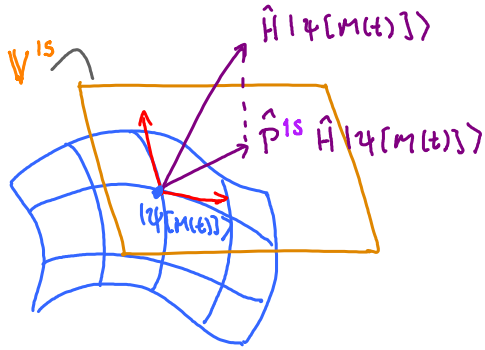


FIG. 4. Kondo-Heisenberg (KH) cylinder: Fermi wave vectors $|k_{Fx}(k_y)|$ for a 40×4 KH cylinder for various values of J_K. Symbols are data points (error bars are below symbol size), lines are guides to the eye. In the insets, black lines sketch the presumed FS for $\mathscr{L}_y \to \infty$, dotted lines show the k_y values allowed for $\mathscr{L}_y = 4$.

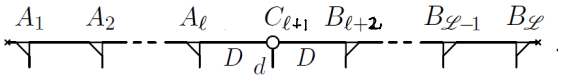
4. CBE-TDVP [Li2022]

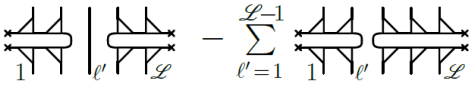
Recall 1s TDVP:

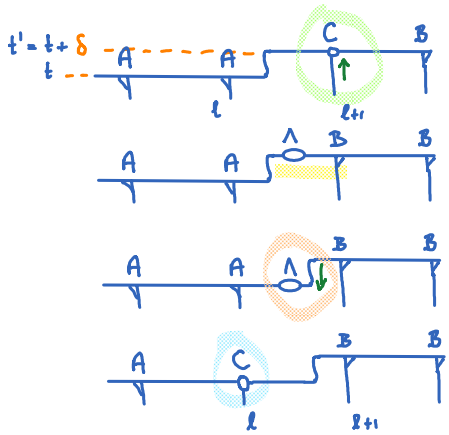


Schrödinger equation for MPS:

$$i\frac{d}{dt}|\Psi[M(t)]\rangle = \hat{P}^{1s}\hat{H}|\Psi[M(t)]\rangle \qquad (1)$$

$\Psi =$  (2)

$\mathcal{P}^{1s} =$  (3)



1s TDVP algorithm (sweeping right-to-left):

(1) Integrate $\quad i\dot{C}_{\ell+1} = H^{1s}_{\ell+1} C_{\ell+1}$ from $t \to t' = t+\delta$

(2) QR factorize $\quad C_{\ell+1}(t') = \Lambda_\ell(t')\, B_{\ell+1}(t')$

(3) Integrate $\quad i\dot{\Lambda}_\ell = -iH^b_\ell \Lambda_\ell$ from $t' \to t$

(4) Update $\quad A_\ell(t)\, C_{\ell+1}(t) \to \underbrace{A_\ell(t)\Lambda_\ell(t)}_{=:\, C_\ell(t)} B_{\ell+1}(t')$ with

Advantages of 1s TDVP: applicable to long-ranged Hamiltonians, numerical stability, unitary time-evolution, energy conservation (because truncation happens before, not after, time step!)

1s-TDVP has two leading errors:

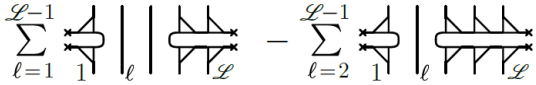
(i) Lie-Trotter error, can be reduced by higher-order integration schemes, e.g. third-order, with error $\mathcal{O}(\delta^3)$

(ii) Projection error, quantified by $\quad \Delta_P = \|(1-\hat{P}^{1s})\hat{H}\psi(t)\|$ (4)

Projection error can be reduced by using 2s TDVP, $\quad i\frac{d}{dt}|\Psi[M(t)]\rangle = \hat{P}^{2s}\hat{H}|\Psi[M(t)]\rangle$ (5)

Then projection error becomes

$\Delta_P = \|(1-\hat{P}^{2s})H\psi(t)\|$

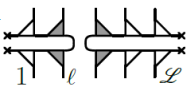
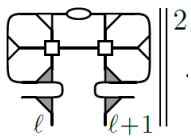
$\mathcal{P}^{2s} =$  (6)

However, after time step, another truncation is needed to bring down bond dimension from Dd to D.

This truncation-after-time-step leads to non-unitary time-evolution, non-conservation of energy.

CBE-TDVP

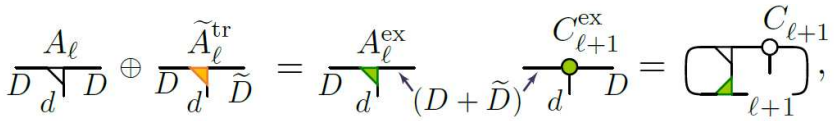
Key idea: use CBE to reduce 2s contribution to Δ_P, given by $\Delta_P^{2\perp} = \|\underbrace{\hat{P}^{2s}(1-\hat{P}^{1s})}_{P2\perp}\hat{H}\psi(t)\|$ (7)

$\mathcal{P}^{2\perp} =$  , $\quad \Delta_P^{2\perp} =$  (8)

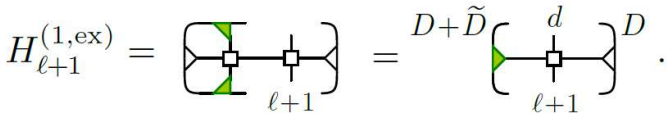
$\Delta_P^{2\perp}$ is the same object is that minimized for CBE-DMRG! Hence, CBE is also useful here!

We add just one step (0) to 1s-TDVP algorithm (when sweeping right-to-left), using:

(0) expand $D \to D + \tilde{D}$ for bond ℓ, using $A_\ell \to A^{ex}_\ell$, $C_{\ell+1} \to C^{exp}_{\ell+1}$, $H^{1s}_{\ell+1} \to H^{1s,exp}_{\ell+1}$,



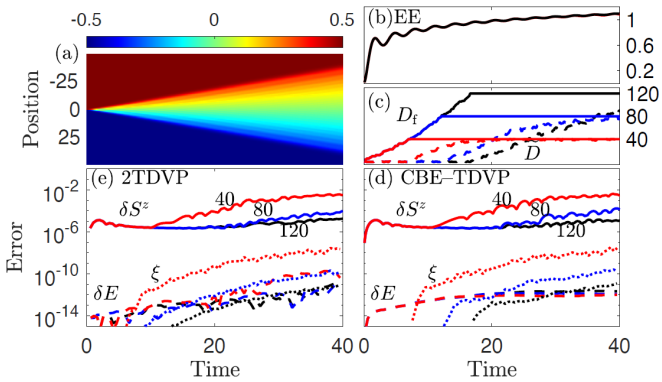
(9)



(10)

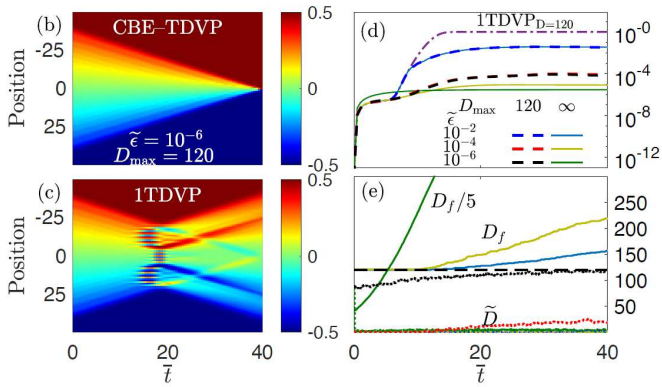
Other steps remain as before, except that in (2), QR factorization is replaced by SVD, to 'trim bond dimension from $D+\widetilde{D}$ to final value D_f, chosen such that truncation error is $< 10^{-12}$. (for early times), or such that $D_f = D_{max}$ (for later times, to limit computational costs). Trimming error is characterized by discarded weight, $\xi(t)$, which can be controlled or monitored. TDVP properties of unitary time evolution and energy conservation hold within $O(\xi(t))$.

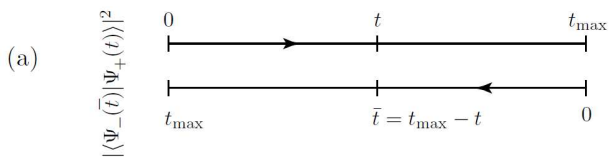
Benchmarking CBE-TDVP for exactly solvable XX model:



$H_{\rm XX} = \sum_\ell (S^x_\ell S^x_{\ell+1} + S^y_\ell S^y_{\ell+1})$ $\quad |\Psi(0)\rangle = |\uparrow\uparrow\ldots\uparrow\downarrow\downarrow\ldots\downarrow\rangle$

FIG. 1. 100-site XX spin chain: Time evolution of a domain wall, computed with time step $\delta = 0.05$ and U(1) spin symmetry. (a) Local magnetization profile $S^z_\ell(t)$. (b) Entanglement entropy EE(t) between the left and the right half of the chain. (c) Bond dimension $D_{\rm f}(t)$ and its pre-trimming expansion $\widetilde{D}(t)$ per time step, for $D_{\rm max} = 120$. (d,e) Error analysis: magnetization $\delta S^z(t)$ (solid line),i.e., the maximum deviation (over ℓ) of $S^z_\ell(t)$ from the exact result, , energy $\delta E(t)$ (dashed line), and discarded weight $\xi(t)$ (dotted line) for $D_{\rm max}=40$ (red), 80 (blue) and 120 (black), computed with (d) CBE–TDVP or (e) 2TDVP. Remarkably, the errors are comparable in size, although CBE–TDVP has much smaller computational costs.

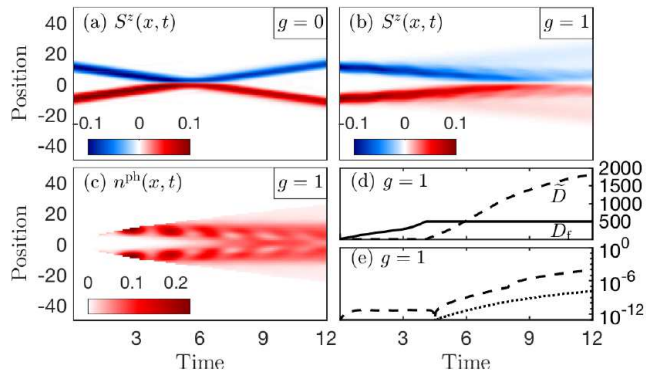




$F(\bar{t}) = |\langle\Psi_-(\bar{t})|\Psi_+(t)\rangle|^2\,, \quad \bar{t}=t_{\rm max}-t \in [0, t_{\rm max}]$

FIG. S-1. (a) Forward-backward time evolution for the computation of $F(t)$. (b,c) Back-evolution of the domain wall, described by $|\Psi_-(\bar{t})\rangle$, computed using (b) CBE–TDVP and (c) 1TDVP. (d) Time evolution of $\delta F(\bar{t}) = 1-F(\bar{t})$, computed via 1TDVP with $D = 120$ (dash-dotted line), and via CBE–TDVP using three values of $\widetilde{\epsilon}$, and either with $D_{\rm max} = 120$ (dashed lines) or $D_{\rm max}=\infty$ (solid lines). (e) Time evolution of the corresponding bond dimensions $D_{\rm f}(\bar{t})$ (solid lines) and $\widetilde{D}(\bar{t})$ (dots). (The solid green curve shows $D_{\rm f}/5$.)

Phonon-induced pair attraction during electron-electron scattering



$$H_{\rm PH} = \sum_\ell U n_{\ell\uparrow} n_{\ell\downarrow} + \sum_\ell \omega_{\rm ph} b^\dagger_\ell b_\ell + \sum_{\ell\sigma} (c^\dagger_{\ell\sigma} c_{\ell+1\sigma} + {\rm h.c.})(-t + b^\dagger_\ell + b_\ell - b^\dagger_{\ell+1} - b_{\ell+1})$$

$n^{max}_{ph} = 8$, $d = 4(n^{max}_{ph}+1) = 36$

FIG. 4. Peierls–Hubbard model: Real-space scattering of two electron wave packets, for $U=10$ and $\omega_{\rm ph}=3$, computed with $\delta = 0.05$, $n^{\rm ph}_{\rm max} = 8$ and U(1) spin symmetry. (a,b) Spin magnetic moment $S^z(x,t)$ for $g=0$ and $g=1$. (c) Phonon density $n^{\rm ph}(x,t)$, (d) bond dimensions, and (e) error analysis: energy $\delta E(t)$ (dashed line) and discarded weight $\xi(t)$ (dotted line), all computed for $g=1$, with $D_{\rm max} = 500$.

(a) Without electron-phonon coupling, two wave packets bounce off each other due to strong U repulsion.

(b) With electron-phone coupling, the wave packets tend to stick together, while (c) phonons get activated.