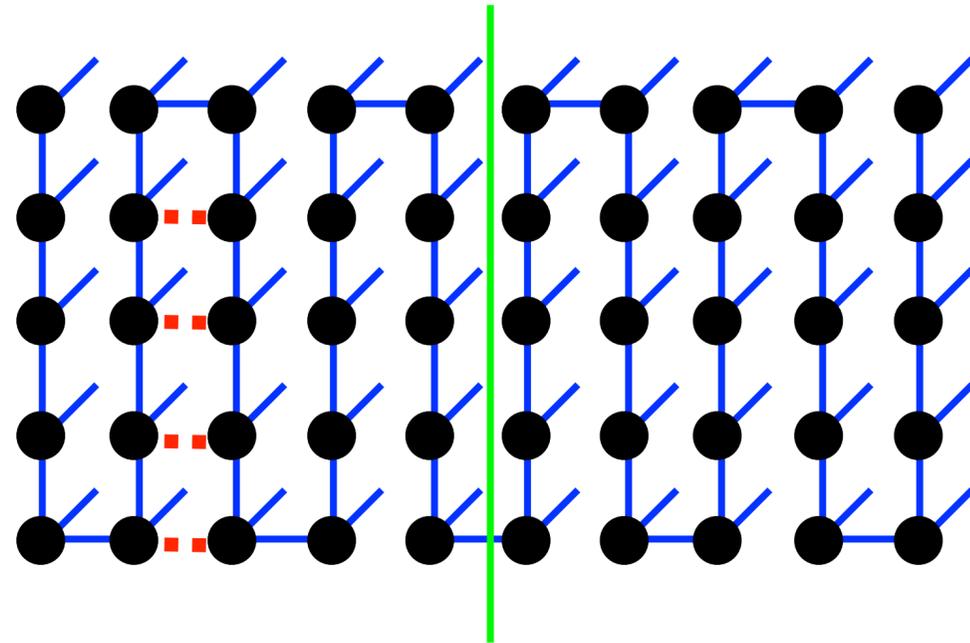
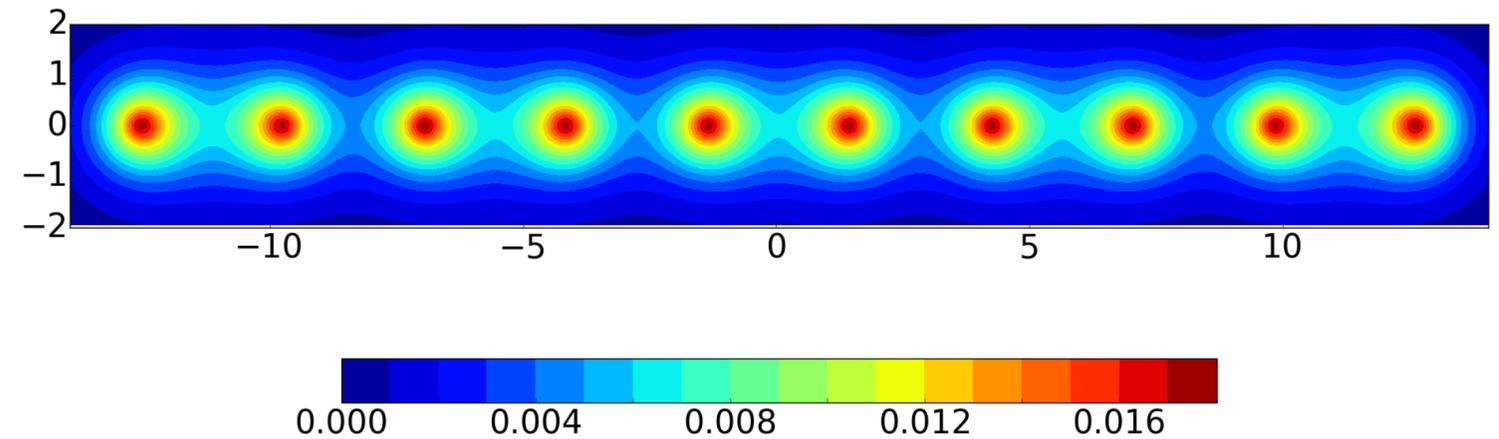


# Model Reduction using localized bases and DMRG

I work primarily in two different areas



DMRG methods for model Hamiltonians describing superconductivity and magnetism

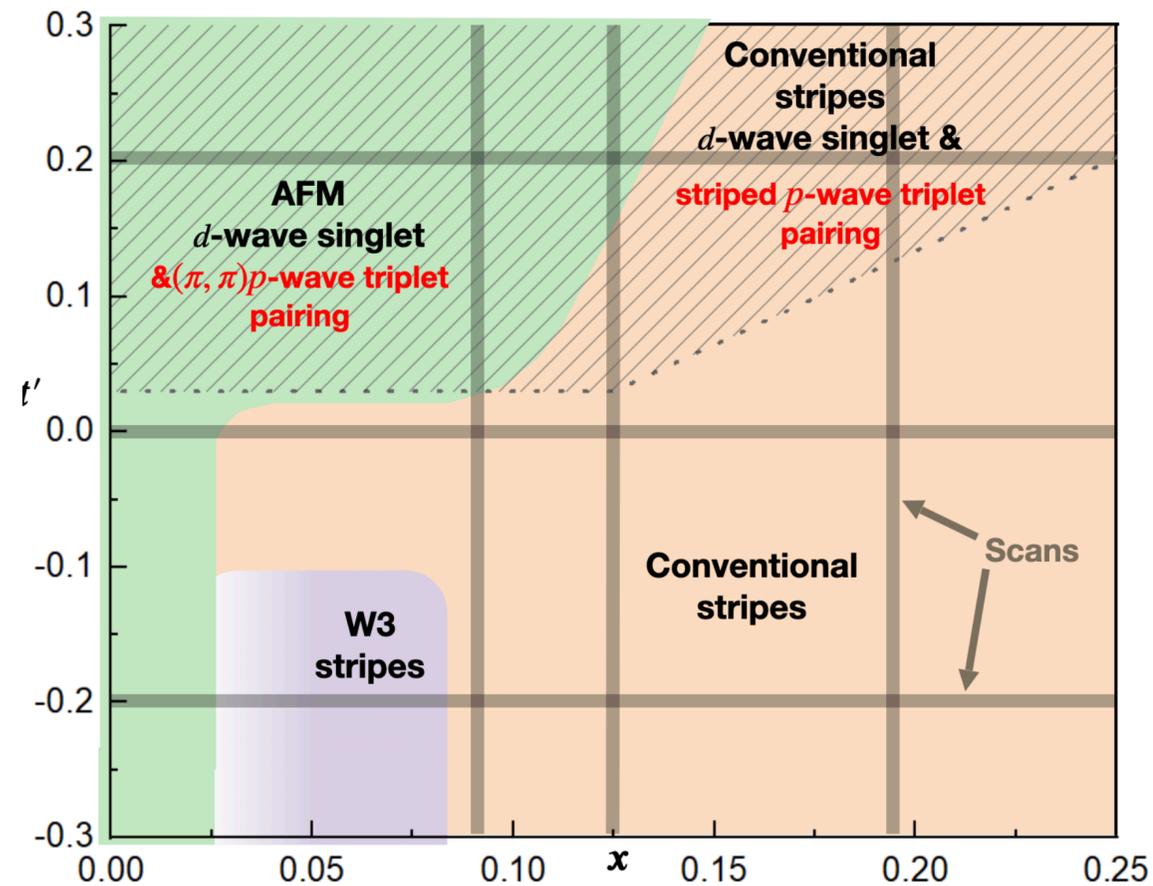


DMRG applied to electronic structure (local bases)

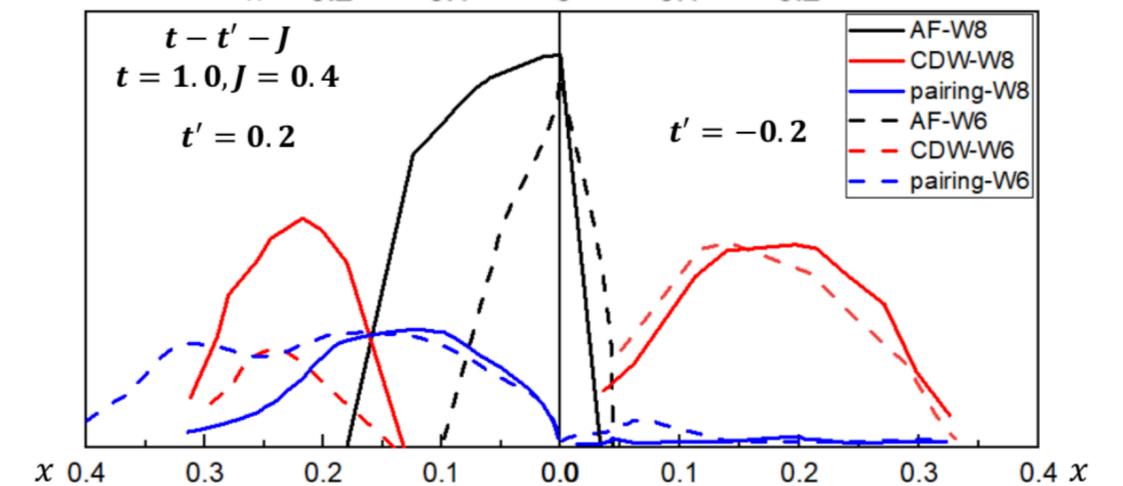
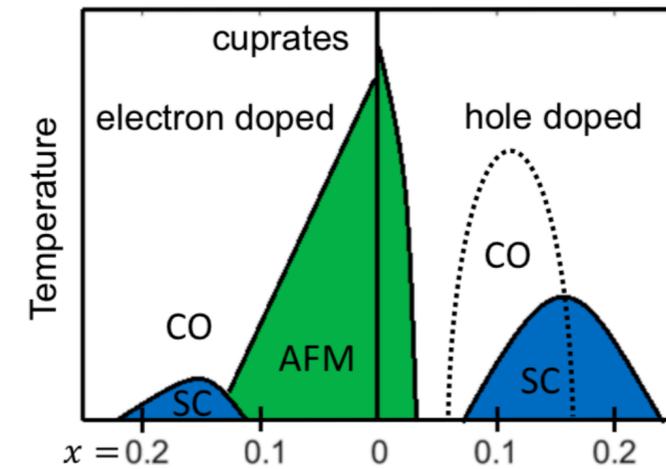
Garnet's talk gave an excellent explanation for why these focuses are linked. This talk is primarily about the second topic, but let me show one slide on the first topic to show the connection

# Phase diagram of $t$ - $t'$ - $J$ model

Jiang, Scalapino, White,  
PNAS 2021



$J/t=0.4$ , based on width-8 cylinders



Compared to the actual materials, we capture many of the properties very nicely — except superconductivity! The hole doped side should have much stronger SC, but we find all the SC is on the electron side. Conclusion: the model needs fixing, but we are not sure how!

# Outline

1. Weaknesses of Gaussian bases for use with DMRG
2. Sliced Bases
3. Derivation of extended Hubbard models for Hydrogen chains using sliced basis DMRG
4. Gausslet Bases

Collaborators: Miles Stoudenmire and Yiheng Qiu (Gausslets)  
Kieron Burke, Tom Baker, Lucas Wagner, Randy Sawaya, Yiheng Qiu, Simons  
Collaboration—sliced basis and related

Thanks to the Simons Foundation, the NSF, and the DOE

DMRG calculations done using ITensor ([itensor.org](http://itensor.org))  
Integral codes and MPO compression in Julia.

# Entanglement and Wavefunction Compression

- A standard way to compress a matrix works if it has low rank, meaning its SVD has a small number of significant singular values  $\lambda$

$$\square = \square \begin{matrix} \diagdown \\ \diagup \end{matrix} \square = \begin{matrix} \square \\ \square \end{matrix} \square$$

- A wavefunction of spins or electrons in Fock space in a basis can be written as a matrix:  $\psi(s_1, s_2, \dots, s_N) = \psi((s_1 \dots s_l), (s_{l+1}, \dots, s_N)) = \psi_{ij}$  do the SVD on this
- The von Neumann entanglement entropy (with respect to the given bipartition) is

$$S_{VN} = - \sum_k \lambda_k^2 \ln \lambda_k^2$$

Small  $S$  means the wavefunction is compressible. The “Area Law” says that for most sensible  $H$ ,  $S$  for the ground state grows with the area of the cut between the two sides rather than either volume. Key ingredient: local  $H$ , also gaps (Hastings)

If you repeatedly cut the system in two in this way, you get a Matrix Product State (MPS) or tensor train. This is the reason DMRG works—DMRG is a low entanglement approximation.

# Quantum chemistry DMRG (White and Martin, 1999, much improved since — esp by Garnet's group)

Key technical issue: there are  $N^4$  interaction terms

$$\frac{1}{2} \sum_{ijkl} V_{ijkl} c_i^\dagger c_j^\dagger c_k c_l$$

- White and Martin: ~25 Gaussians, transformed basis to HF molecular orbitals, to treat one stretched water molecule
- Progress in two decades: ~100 active orbitals, selected from bigger Gaussian basis

This is in contrast to Hubbard-like models, where calculation time scales as  $N$ , so  $N \sim 10^3$  is fine for a chain

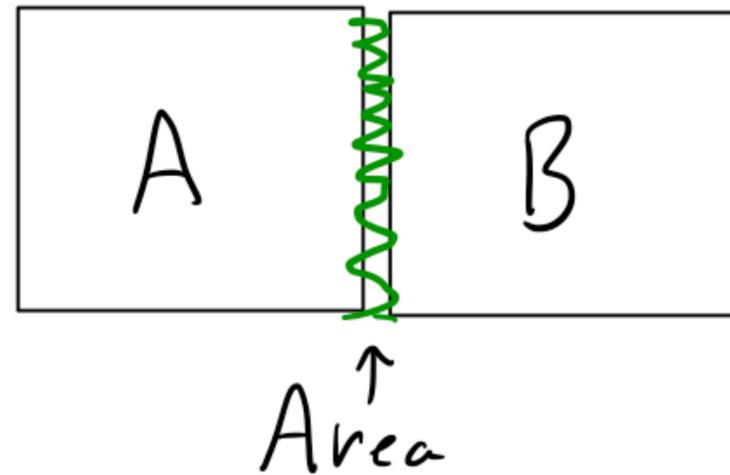
Gaussians are highly optimized to reduce  $N$ —not the number of Ham terms, nor the entanglement. For example, simple grids have  $N^2$  interactions

$$\frac{1}{2} \sum_{ij} V_{ij} n_i n_j$$

Much larger  $N$

# Area Law and continuum problems in a basis

- The Area law is the key to the success of DMRG in lattice models (spin)



$$S \propto A$$

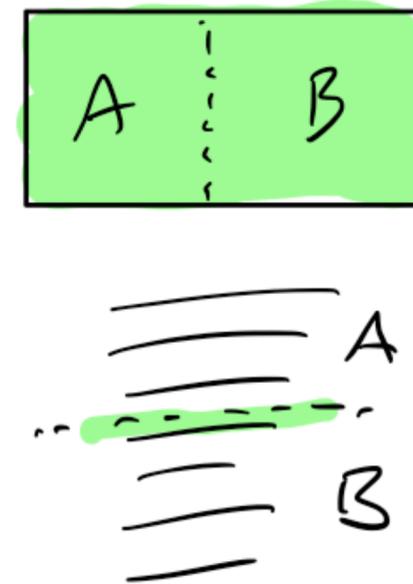
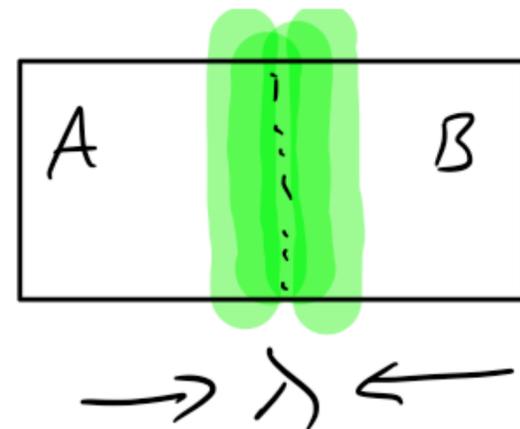
DMRG: matrix size  $m \sim e^{\alpha S}$



- What about in a basis? k-space has no locality,  $S \propto V$
- Exception: noninteracting/Hartree Fock,  $S=0$  if basis=HF orbs

Mostly localized basis:  
Expect: (theory?)

$$S \sim A\lambda$$



Core orbs are well treated by HF, so we might want to switch from HF orbs for some degrees of freedom to localized orbs for others...

# What is the best basis to use with DMRG?

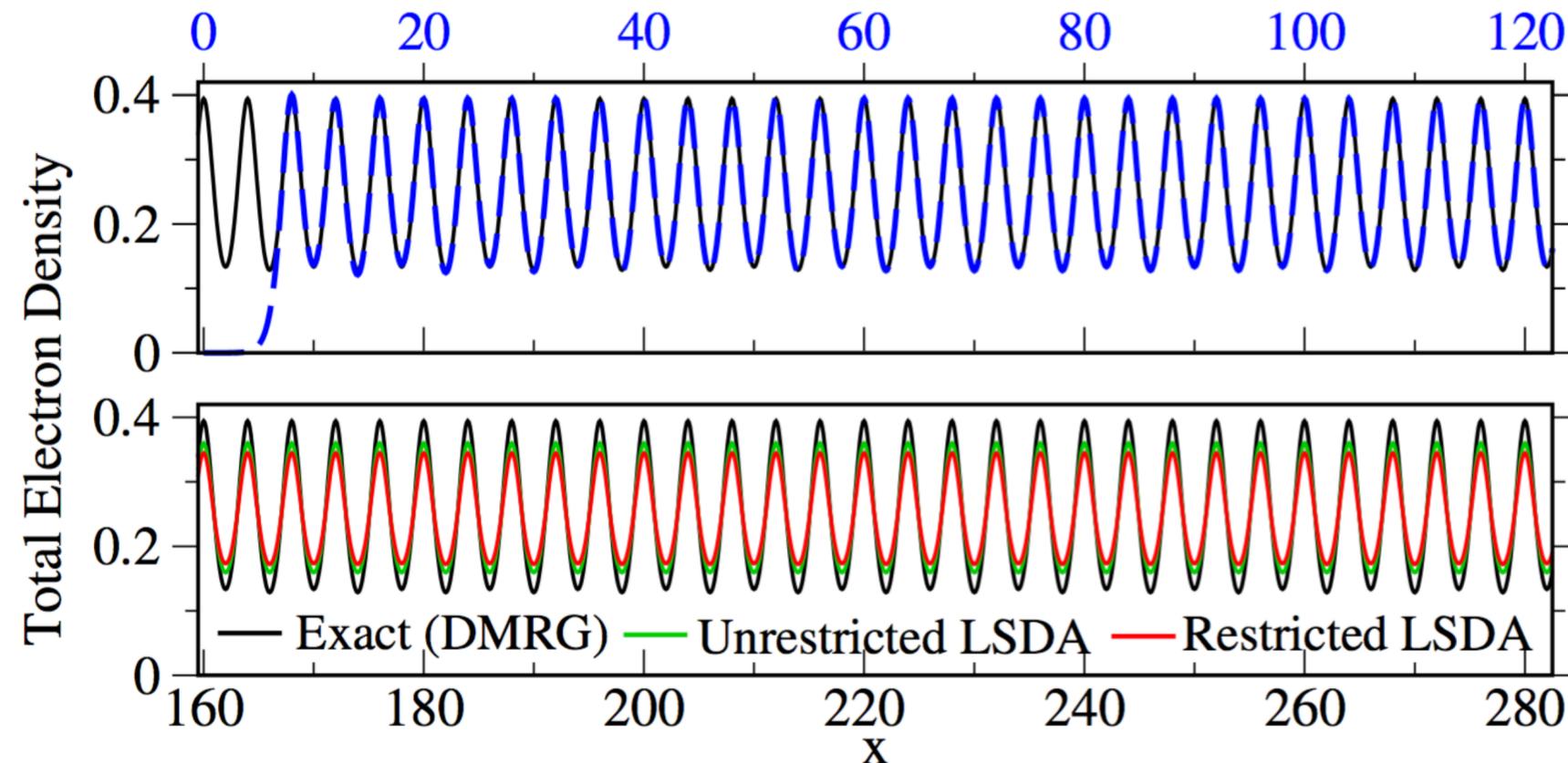
- Area Law (localized in real space) versus HF determinant (energy-localized)
- An approach developed by Garnet's group: localize separately within "occupied basis space" and "unoccupied space". Alternatively, Legeza has developed on-the-fly basis rotations to reduce entanglement. Both much better than pure HF orbs.

Computational issues: any standard basis gives a two electron integral  $V_{ijkl}$  which greatly worsens computational scaling.

Why not use a grid, which gives a diagonal  $\hat{V} = 1/2 \sum_{ij} V_{ij} n_i n_j$  even if the number of grid points is larger?

# Grid based DMRG, 1D

- Along with Kieron Burke's group we developed 1D grids and associated DMRG for 1D continuum models, with soft Coulomb or exponential interactions, to understand DFT
- Grid gives interactions as the favorable  $V_{ij}$ , which we further compressed into a matrix product operator: Result:  $N^2 \rightarrow N$ , linear scaling!
  - Up to 100 pseudoatoms—4000 lattice sites, essentially exact results
- This is a great approach to 1D continuum problems!



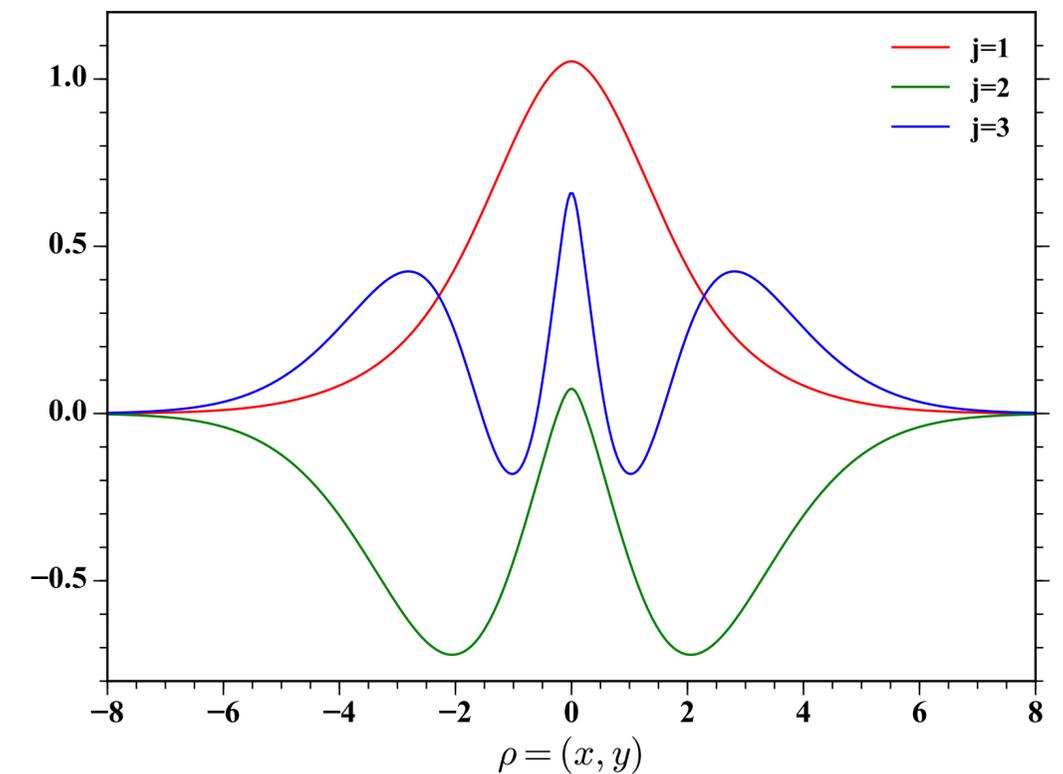
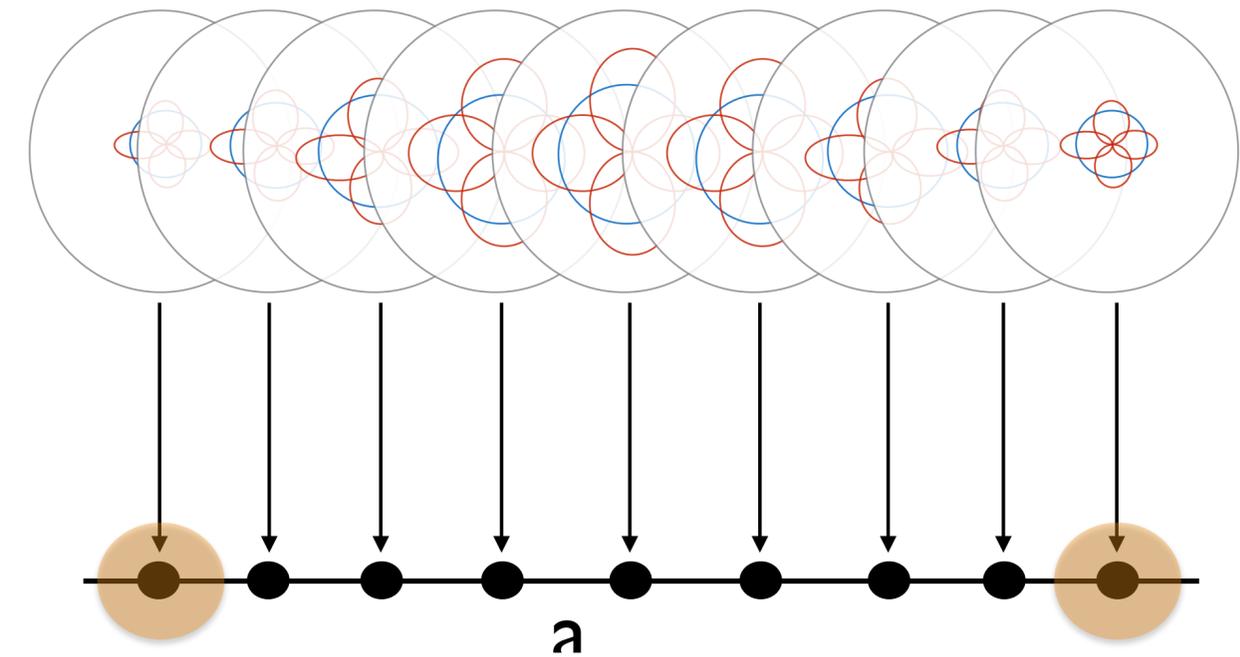
Miles Stoudenmire, Kieron  
Burke, Lucas Wagner, ...

# Grid representations for 3D Electronic Structure?

- A naive 3D grid would give millions of points, but what about a grid in only one direction??

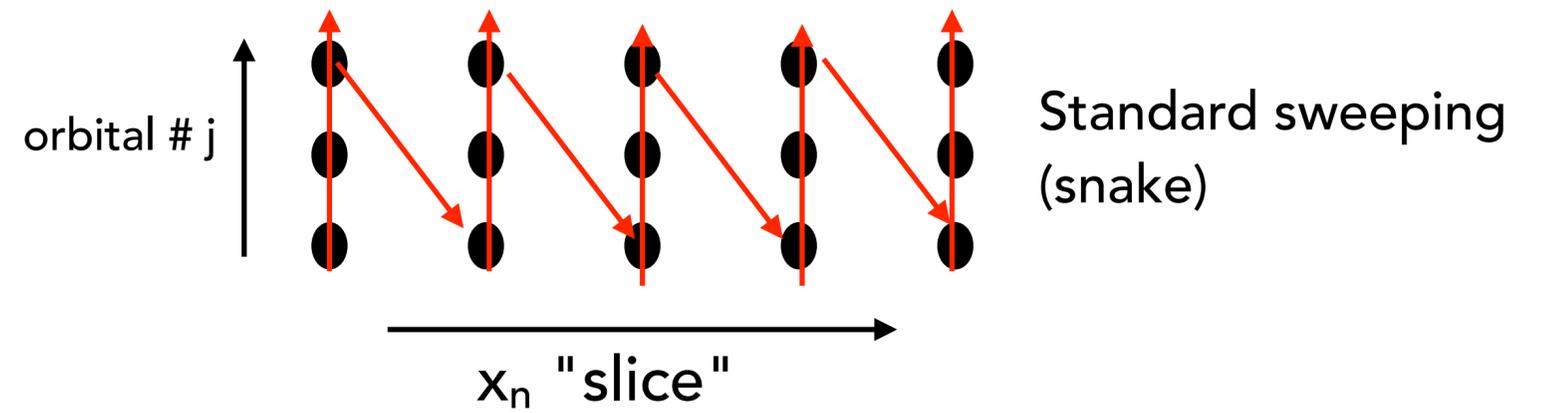
This is just as good as a 3D grid in terms of entanglement, and, it gives block sparsity for  $V_{ijkl}$

In **Sliced Basis DMRG** (Stoudenmire & SRV) we use a grid in the  $z$  direction, 2D Gaussians in  $x$  and  $y$  directions. The Gaussians are slices of standard Gaussian basis sets

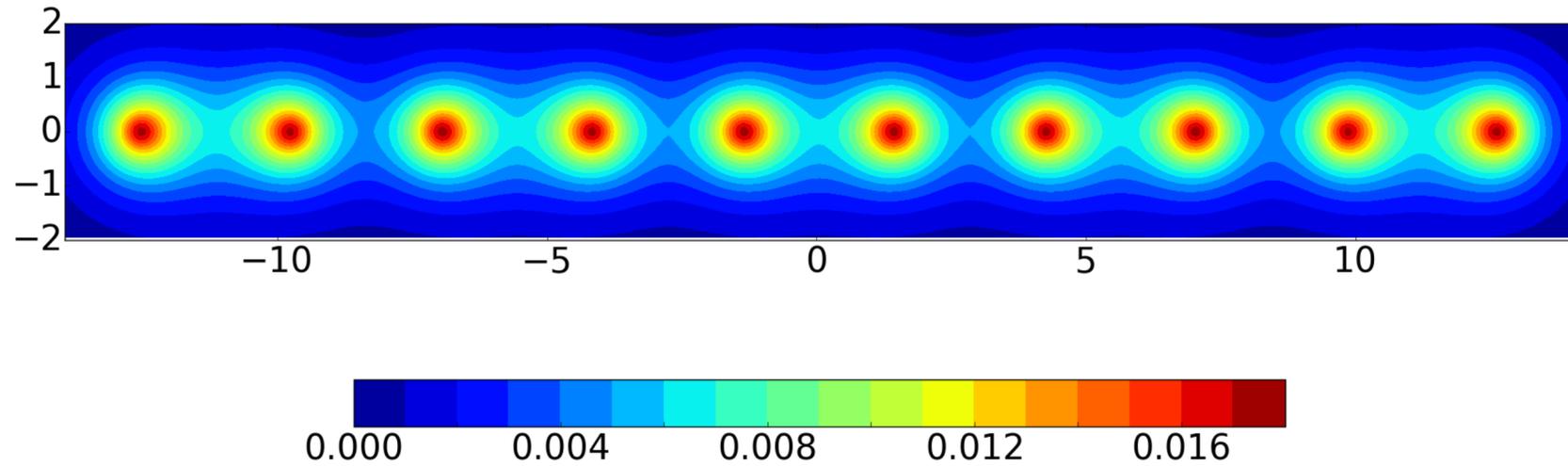


Transverse S functions

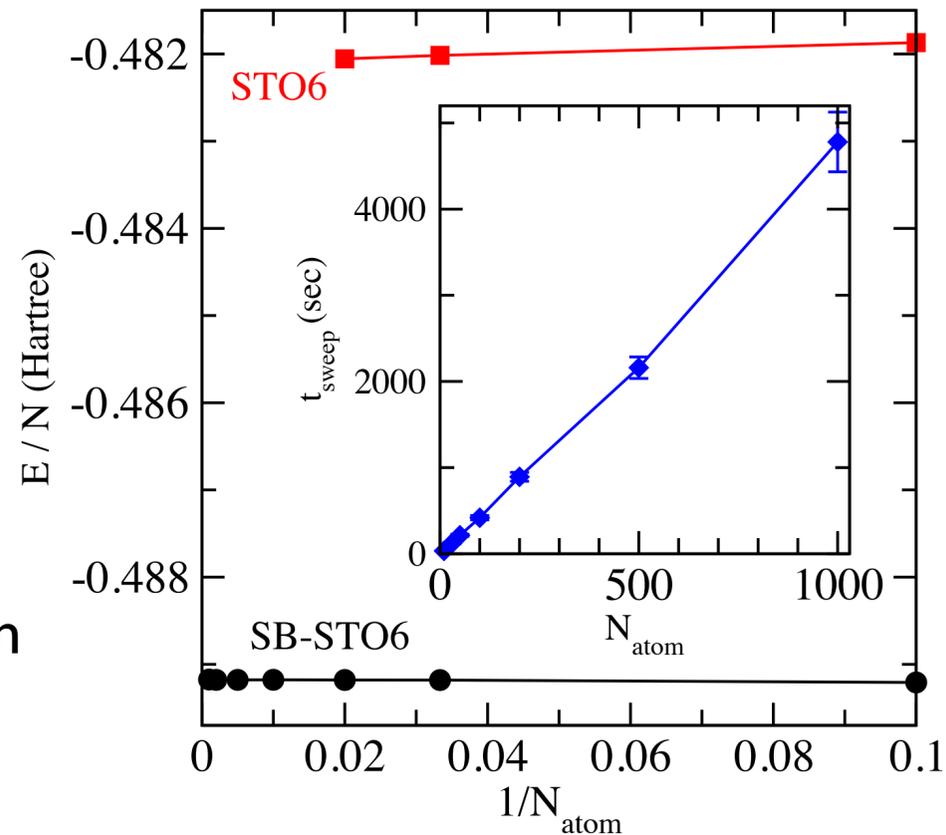
To DMRG, a sliced basis looks like a long multi-leg ladder Hubbard model, and our standard software (ITensor) can be used with little change.



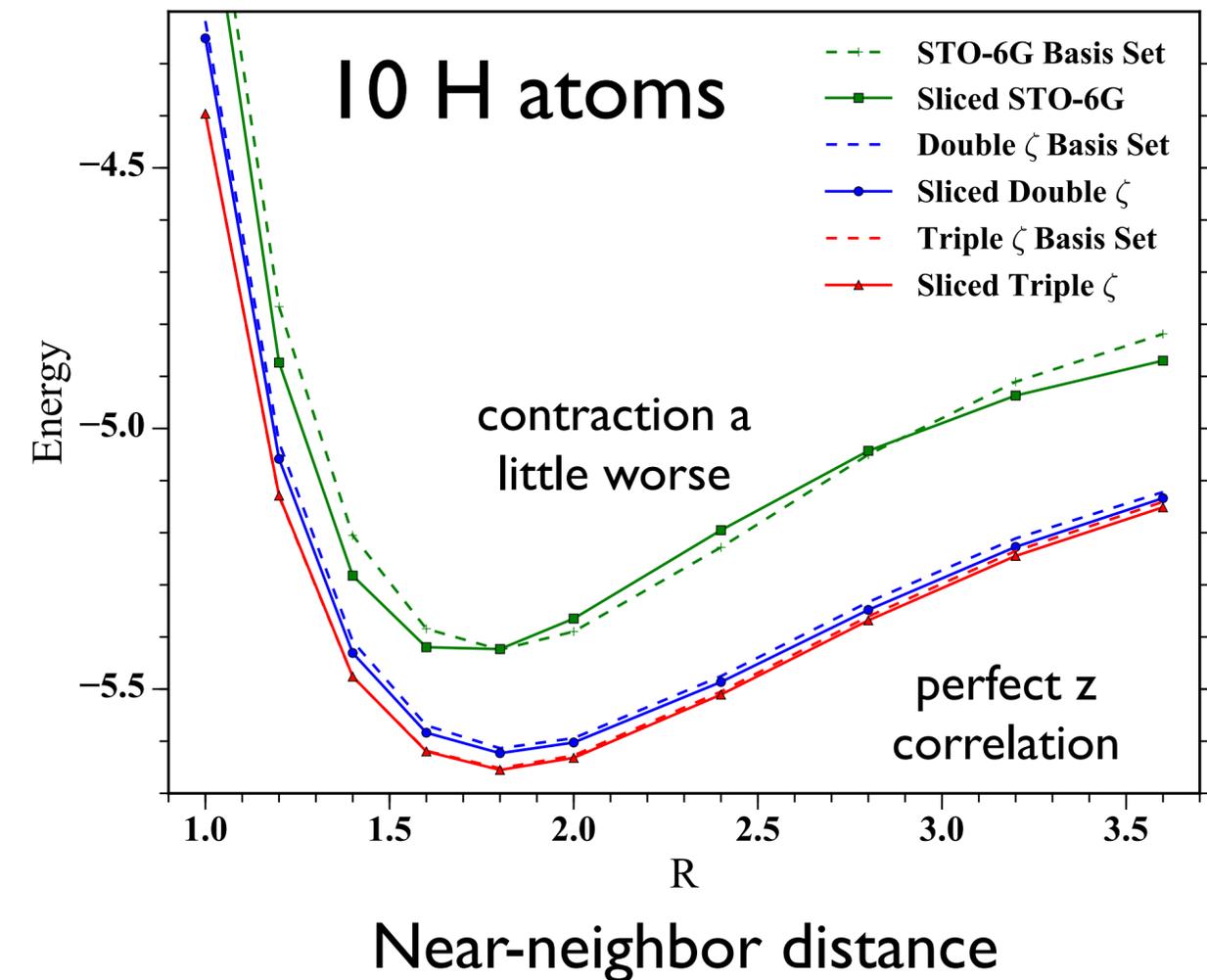
## Application: Hydrogen Chains



Computation time:  
Linear scaling, starting at small  $N$ ;  
1000 atoms on a desktop,  
sub-mH accuracy correlation  
within a minimal basis



## Comparison with QCDMRG



# Model reduction: deriving Hubbard models for the H Chain

- Typical model derivations start with DFT, which is inaccurate for a stretched H-chain. What if we had the exact H-chain solution: how would we derive an effective model?
- Wannier functions: instead of DFT or HF occupied orbitals, use the lowest Natural orbitals to define Wannier functions (Koch&Boedeker) While truncation to  $N/2$  NOs is terrible, truncation to  $N$  omits only about  $10^{-3}$  in the occupancy. [State-averaged RDM too]
- Defining the Wannier functions: localize the  $N$  NOs by diagonalizing the z-coordinate matrix  $\langle j | z | j' \rangle$
- In comparison with a conventional HF approach, the Wannier functions are almost identical (even though HF gives a poor ground state)! Reassuring to conventional approaches.
- Generates fairly short range hoppings (evaluate integrals with WFs).
- Coulomb integrals: the very low occupancy of the omitted NOs tells us they can't screen much. Verification: quantitative agreement between the original sliced basis and WF basis.
- Truncation to  $V_{ij}$  interaction: a simple truncation was very accurate

Randy Sawaya and SRW, PRB

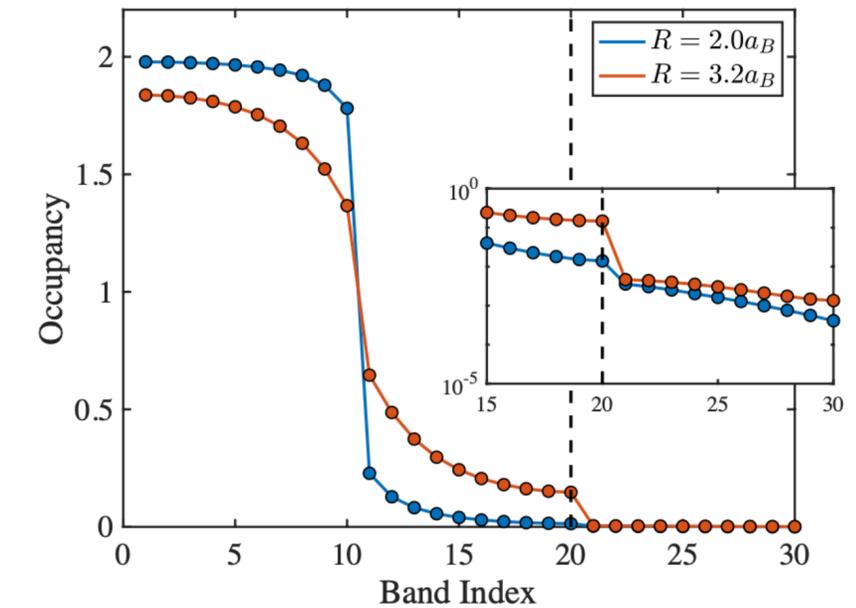
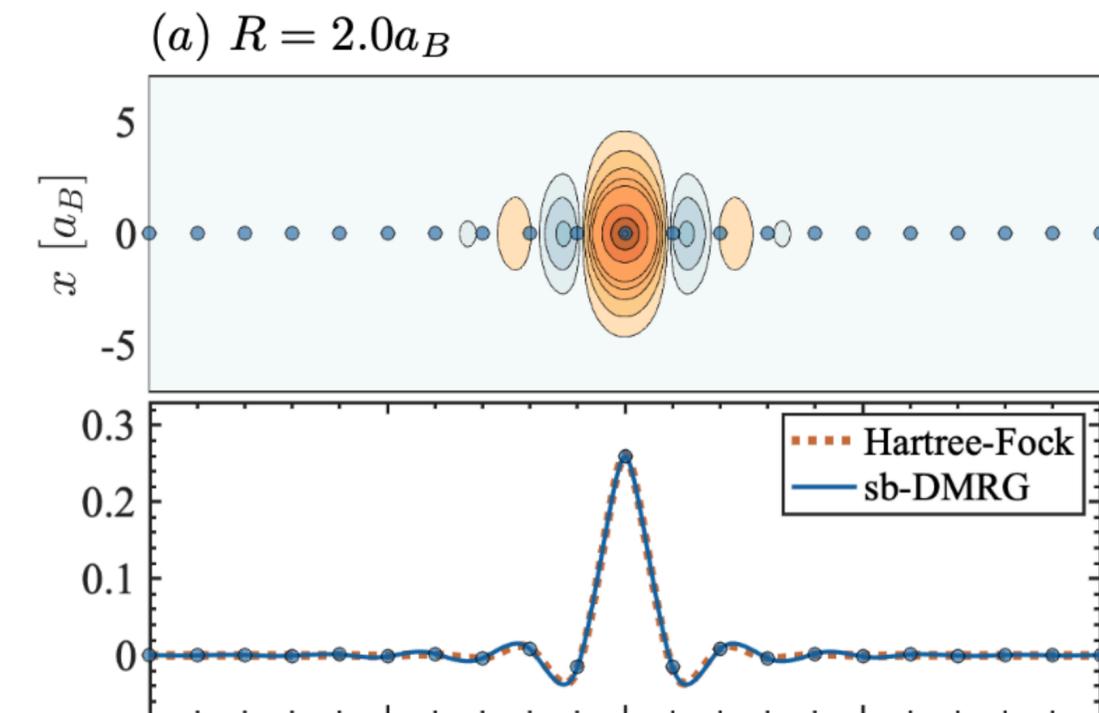


FIG. 1. Eigenvalues of the single-particle correlation matrix from the ground state of H<sub>20</sub> giving the natural orbital occupancy. The fit



# Deriving Hubbard models for the H Chain

- “Strong” model reduction:
  - Note that the half-filled Hubbard model is natural for the H-chain (at larger  $R$ —interesting diffuse band occupancy at small  $R$ )
  - The Mott-Hubbard gap should make the long-range Coulomb interaction “removable” (?)

General strong reduction approach: define the suitability of a Hamiltonian by the low lying states it generates, e.g.  $|\psi_{gr}(H_{eff})\rangle$ , and minimize:  $\langle \psi_{gr}(H_{eff}) | H_{full} | \psi_{gr}(H_{eff}) \rangle$ .

This assumes the two H’s live in the same space, or we have a map between them (Lucas Wagner; Schuler et al PRL 2013)

Reducing all the way to a Hubbard model is OK (with the right parameters) but for very high accuracy, it is better to extend the Coulomb interaction a few atoms. The onsite Hubbard term is completely different if you extend the interaction.

Final checks: the models do an excellent job at long-range correlations and spin excitation energies.

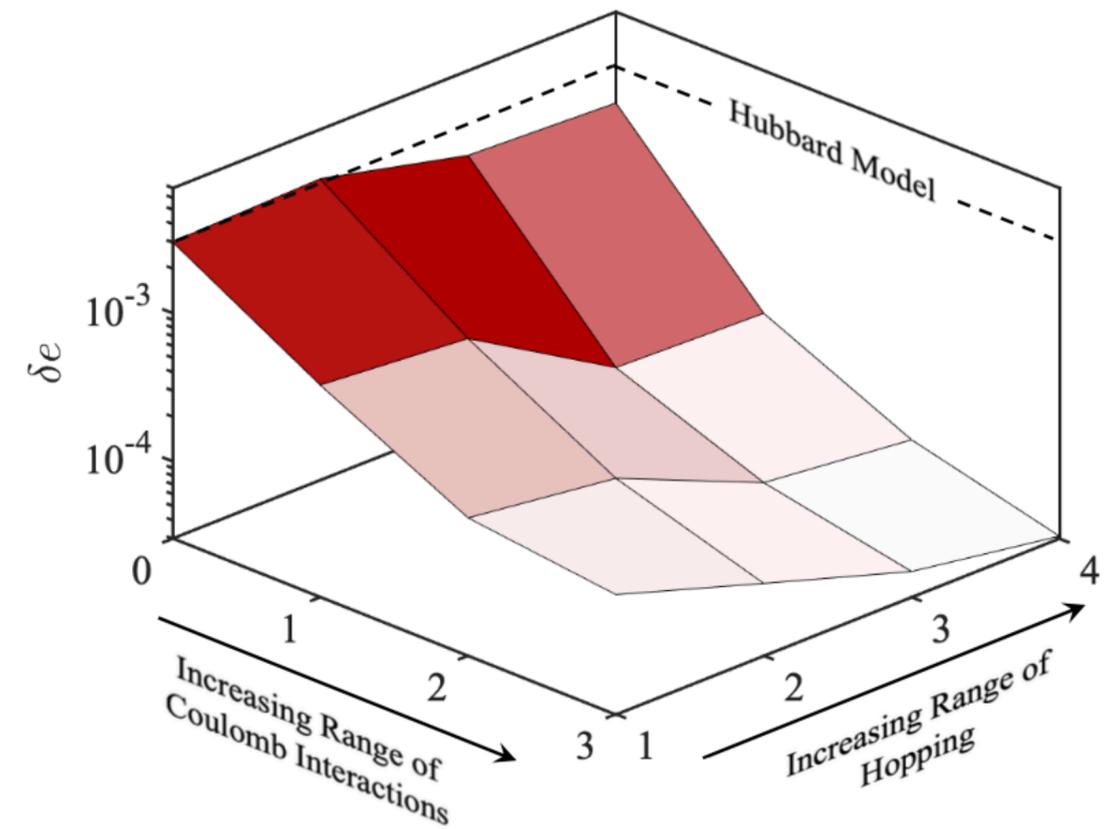
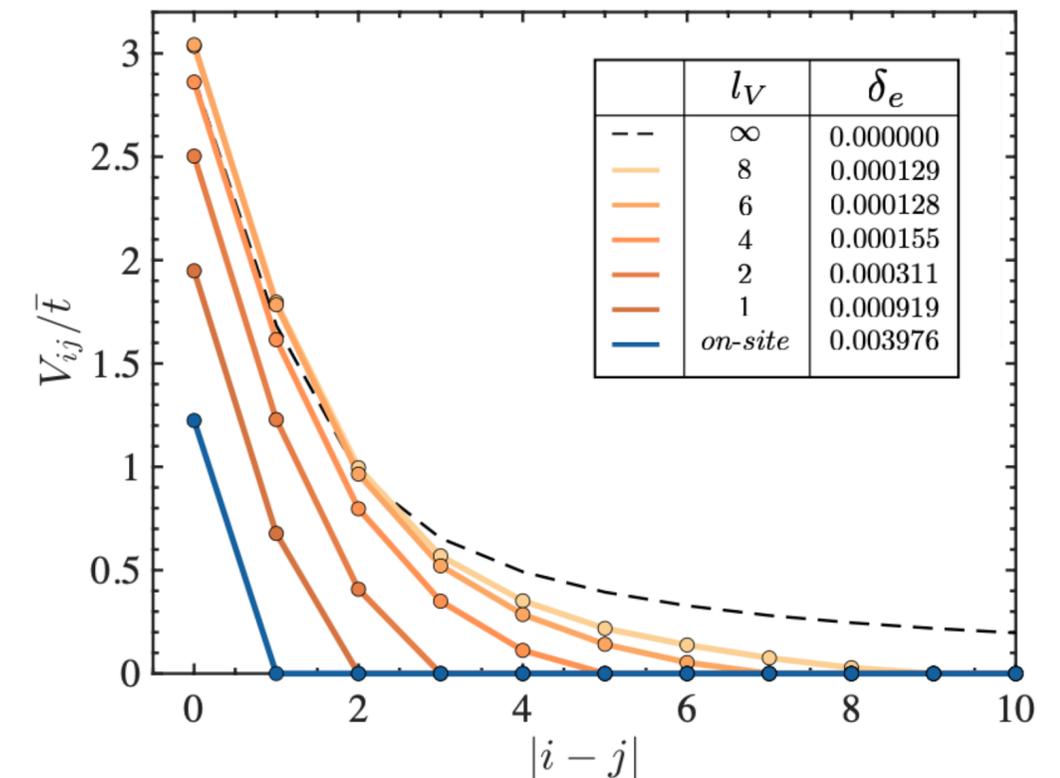


FIG. 7. Error for effective models with varying-range one-body and two-body interactions for H20 at  $R = 2.0a_B$ . The dashed black lines



# Going beyond sliced basis sets: Factoring, DVR, etc?

- Sliced basis have poor scaling in the number of transverse functions  $M$ ,  
 $M \lesssim 15$ ,  $N_{sl}^2 M^4$  terms
- Need a more compact representation. What about factoring  $V_{ijkl}$  ?
  - A factored relationship disappears once you make a compressed MPO, so we don't know how to make use of factoring in DMRG.
- Discrete variable representations (DVR): gives a diagonal interaction! (e.g. sinc basis) But, DVRs seem too delocalized for DMRG, worry about volume law entanglement

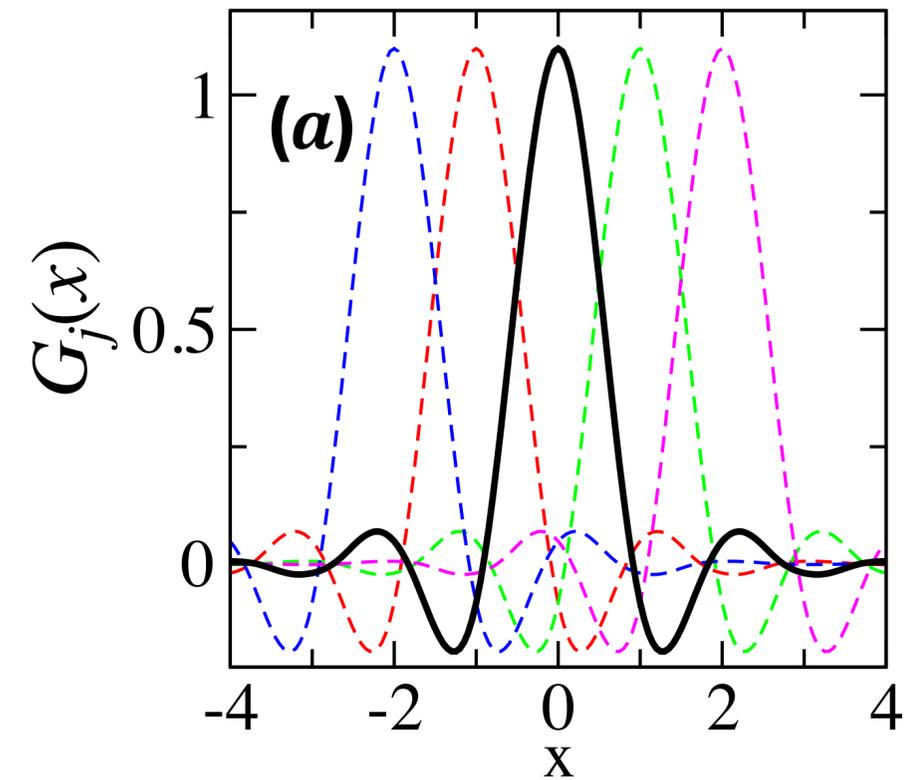
Wavelets were designed for achieving locality, completeness and orthogonality. Can they help?

# Wavelet inspired approach

- Let's go back to a grid: suppose we could use a very coarse grid, in fact with variable spacing (small at nucleus), so that only a modest number of grid points were needed, but we still had the diagonal  $V_{ij}$  form for the el-el term?? This would be *ideal* for DMRG or tensor networks, and perhaps many other approaches.
- The first step: instead of an ordinary grid, use a grid of special local basis functions. Our wish list for these functions:
  - Minimal width (“uncertainty”) in  $x$  and  $p^2$
  - Orthonormal
  - Complete (exact fit to any polynomial up to a specific order)
  - Integrates like a delta function up to some polynomial order

$$\int dx G(x)x^n = \delta_{n,0}$$

Remarkably, if our functions satisfy these requirements, then they give a diagonal representation,  $V_{ijkl} \rightarrow V_{ij}$  as an approximation which converges rapidly with decreasing grid spacing



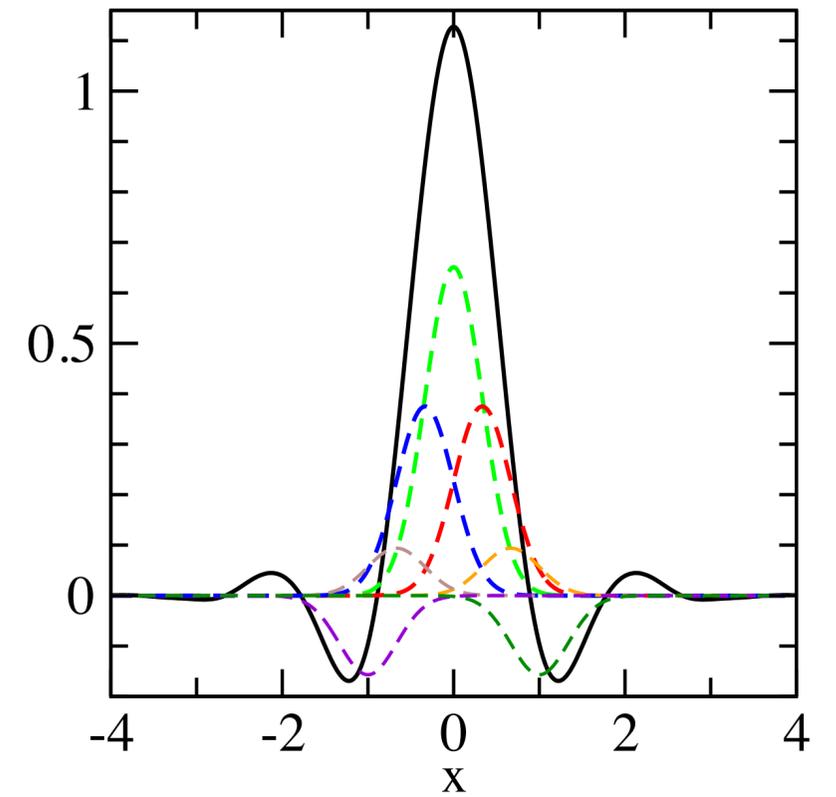
“Gausslet”

Trying to get the best properties of a grid and basis at the same time!

## Diagonal Approximations from the delta-function property

Assume  $\int d^3r S_j(\vec{r})f(\vec{r}) = f(\vec{r}_j)$  where  $\vec{r}_j$  is the center of basis function  $S_j$ , for smooth functions  $f$ . Then the expansion coefficient (orthonormality) of  $f$  at point  $j$  is also  $f(\vec{r}_j)$ . Given a potential  $V(\vec{r})$  and wavefunction  $\psi(\vec{r})$ , the expansion coefficient of  $V\psi$  is  $V(\vec{r}_j)\psi(\vec{r}_j) = V_j\psi_j$ . Thus the diagonal matrix  $V_j$  acts as the potential energy matrix.

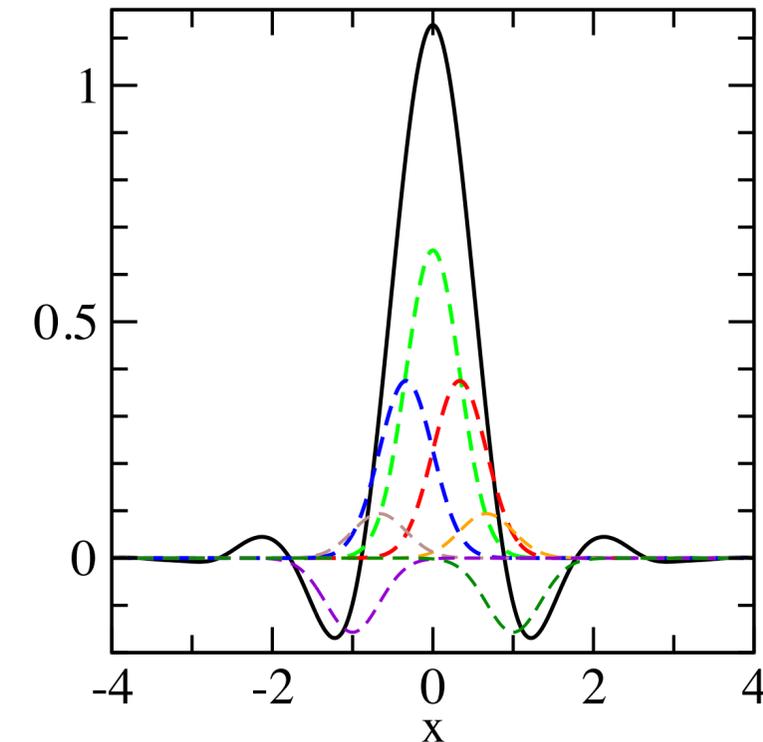
The same property holds for the two electron interaction,  
 $V_{ijkl} \rightarrow V_{ij}$



“Gausslet”

# Gausslets

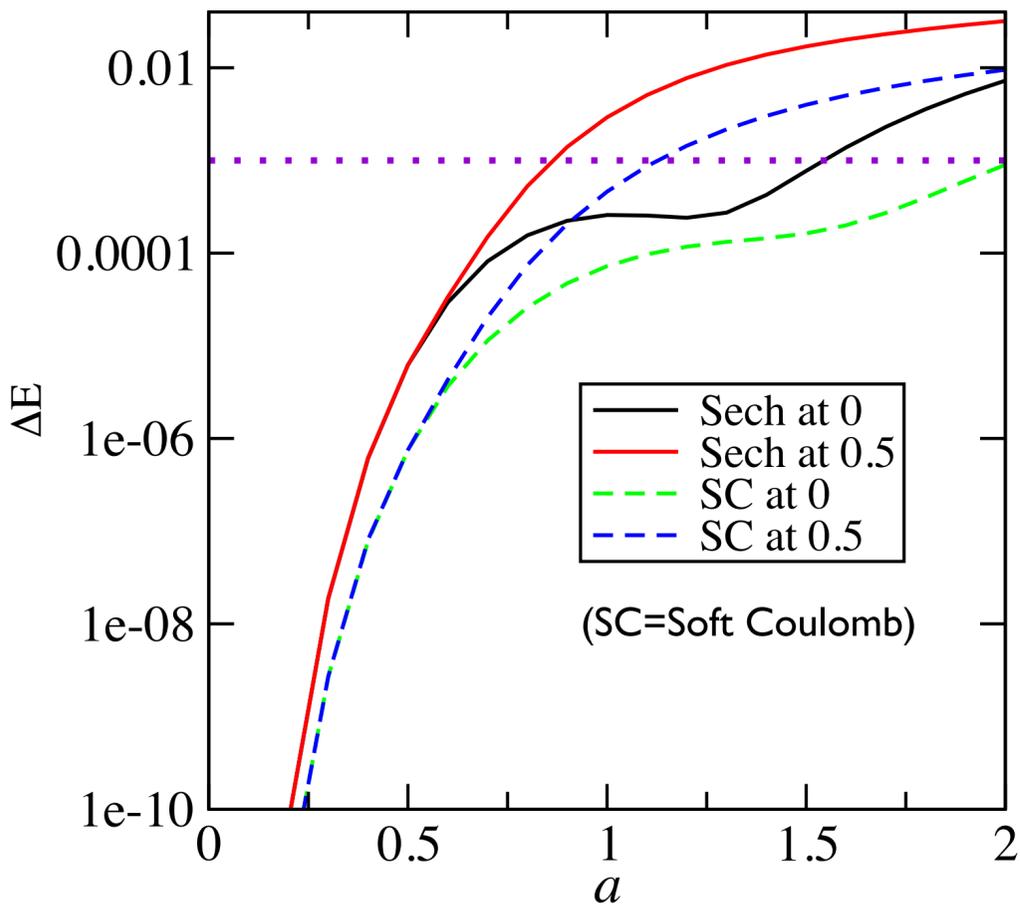
- How do we find local, orthogonal, complete functions satisfying the delta-function property?
- Standard orthogonal wavelet-scaling functions (Daubechies) satisfy compactness, orthogonality, completeness
- One specific set(coiflets) also have the delta-function property.
- These standard functions are not symmetric, with no closed form for the functions. Integrals require unfamiliar wavelet tricks
- What are gausslets? Glen Evenbly and I used a close relationship between wavelets and tensor networks (MERA) to develop symmetric ternary wavelets. These are better than standard wavelets (see Glen's IPAM talk).
- Gausslets are based on those, but I was able to construct them as a contraction over a grid of constant-width Gaussians, for easy integrals and a simple analytic closed form.
- 3D:  $G(x)G(y)G(z)$ , which is a sum of 3D Gaussians: integrals are easy. Easy to combine with standard Gaussian bases



“Gausslet”

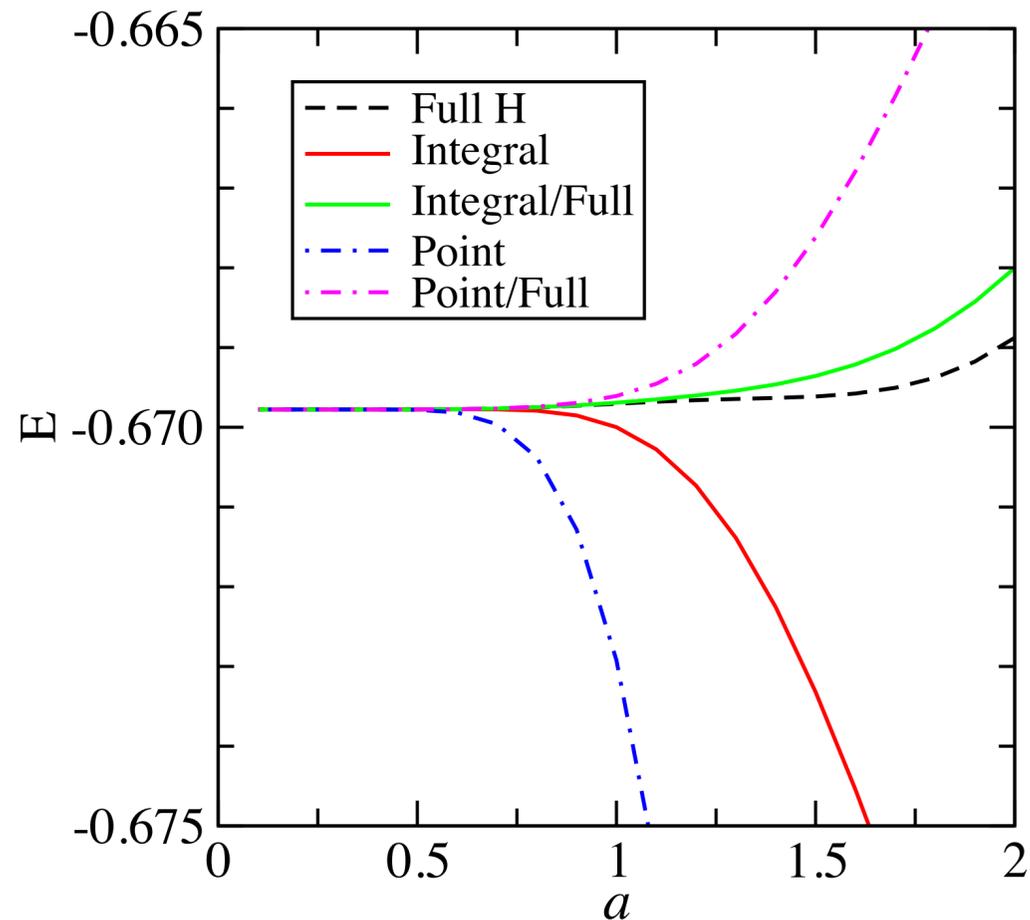
$$G(x) = \sum_j b_j \exp\left[-\frac{1}{2}(3x - j)^2\right]$$

# 1D tests of Gausslets

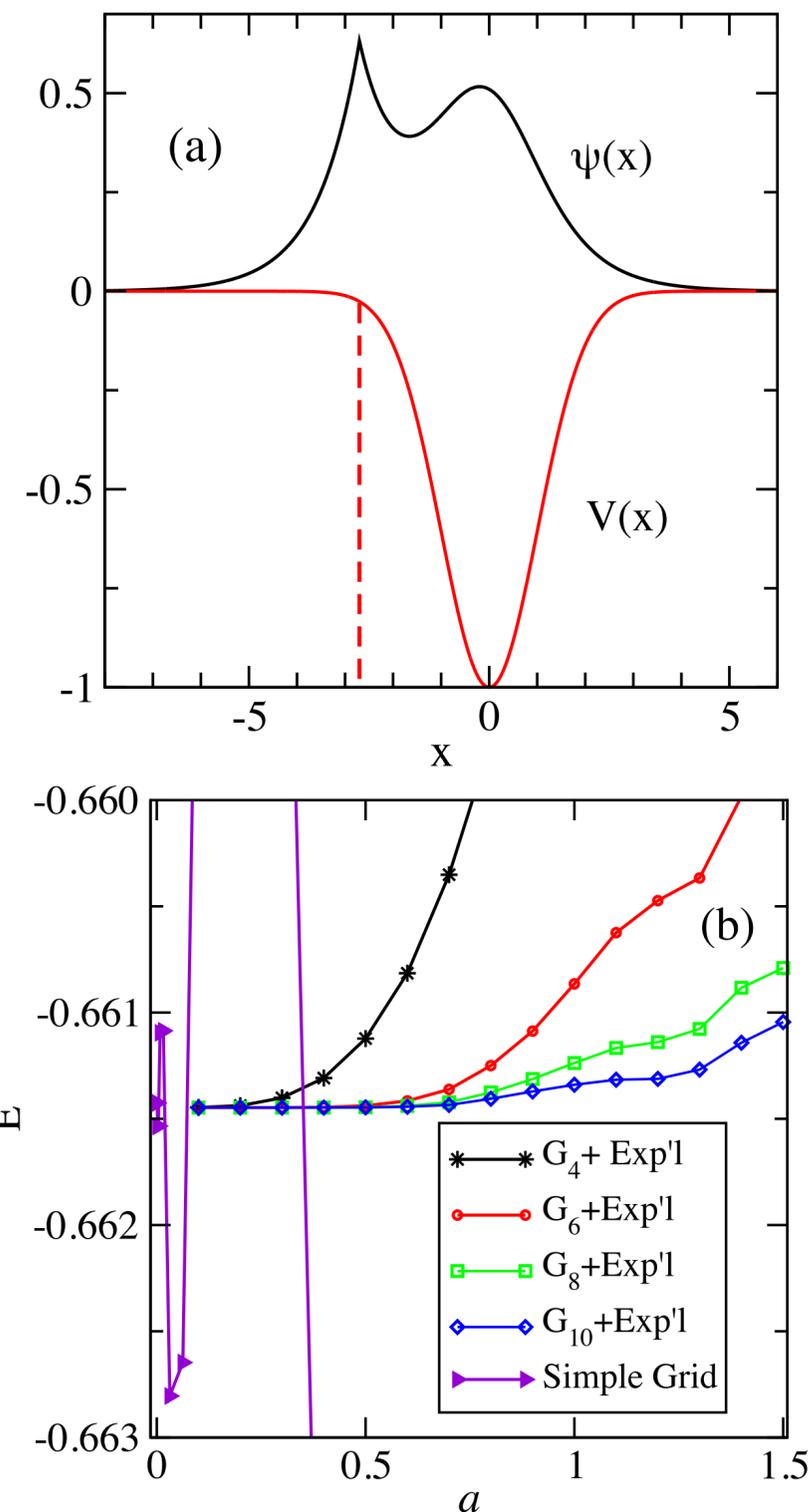


Convergence of gausslet basis  
Noninteracting potentials

$a$  = spacing of the gausslets



Tests of diagonal approximations  
Integral/Full and Point/Full are  
using a diagonal approx to get the  
ground state, and then plugging  
into the exact H



Why are a grid of gausslets better  
than a grid? You can add extra  
basis fns adapted to a singularity.

So far this makes a great 1D basis—need a few  
more improvements to make 3D really practical

# Variable Resolution—more basis fns near nuclei

- How can we get resolution which varies in space while maintaining orthogonality and diagonal approximations?
- Coordinate transformation: Let  $u \rightarrow x(u)$  be a smooth, one-to-one function. Let  $S_i(x)$  be a set of orthonormal gausslets. Define a distorted gausslet

$$\tilde{S}_i(x) = S_i(u(x)) \sqrt{u'(x)}$$

Then

$$\int_x \tilde{S}_i(x) \tilde{S}_j(x) = \int_u [u'(x)]^{-1} \tilde{S}_i(x(u)) \tilde{S}_j(x(u)) = \int_u S_i(u) S_j(u) = \delta_{ij}$$

If  $S_i(u)$  integrates like a  $\delta$ -function, so does  $\tilde{S}_i(x)$

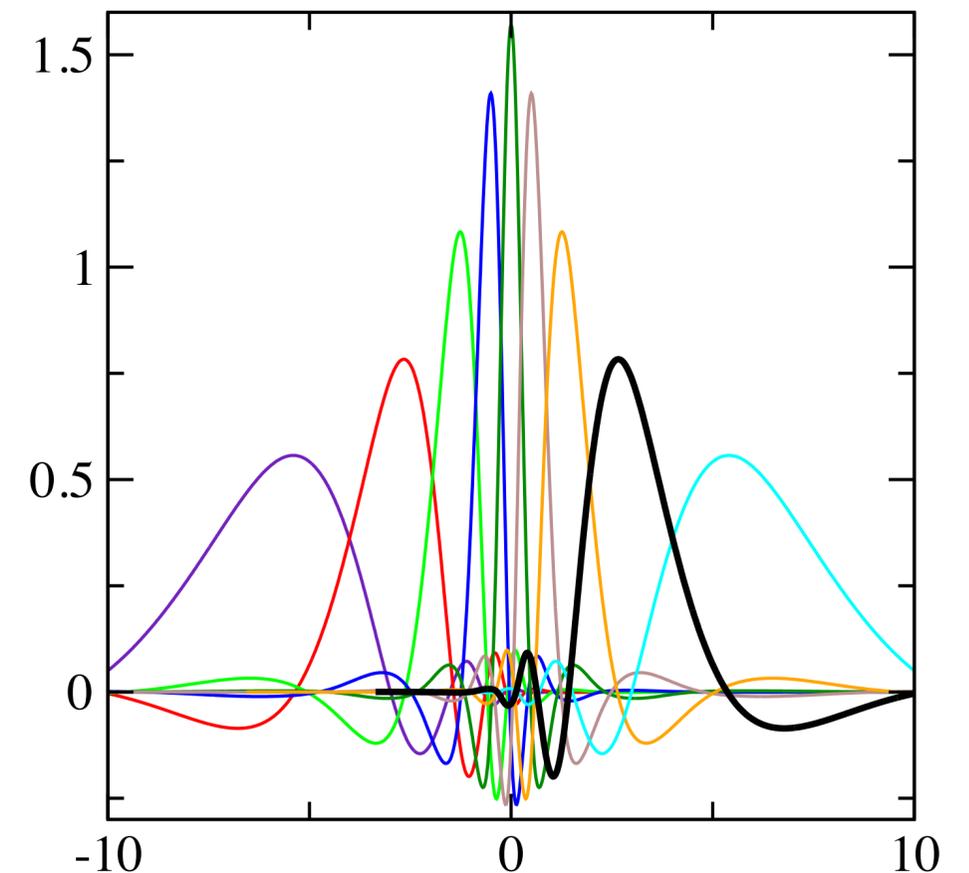
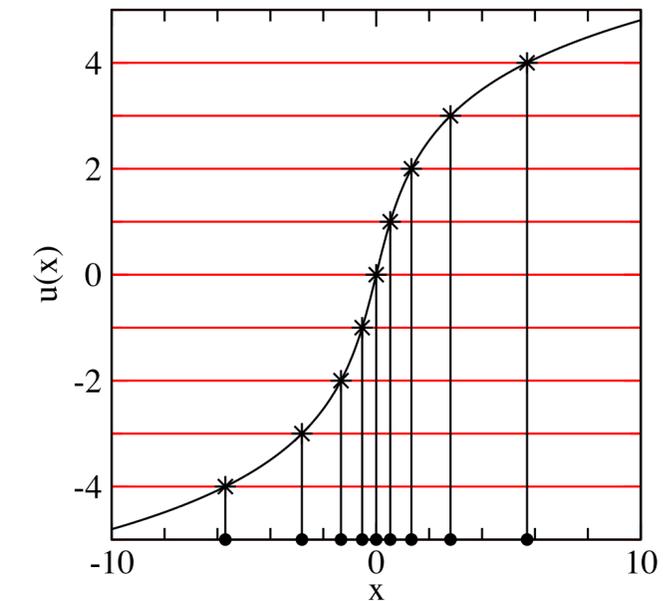
$$\int_x \tilde{S}_i(x) P(x) = \int_u [u'(x)]^{-1} \tilde{S}_i(x(u)) P(x(u)) = [x'(u_i)]^{1/2} P(x(u_i))$$

For this to be effective, the mapping must be very smooth. A simple analysis suggests the “inverse sinh” transformation

# Single atom transformation

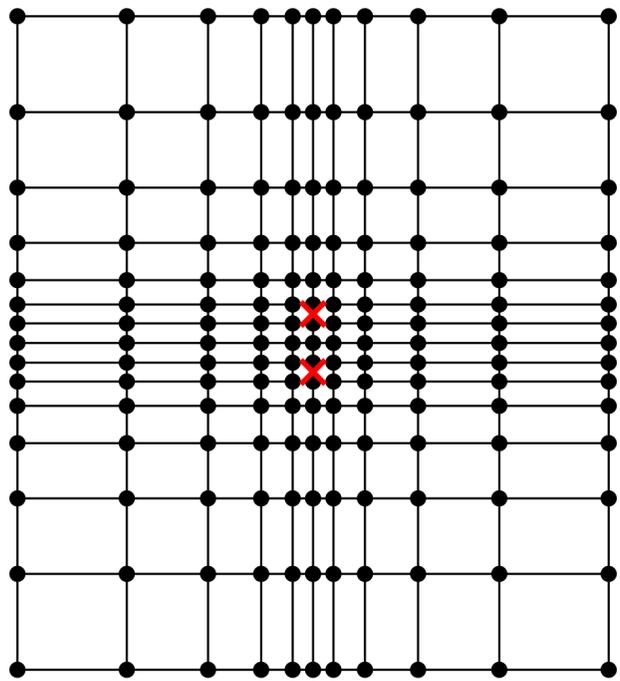
$$u(x) = \frac{1}{s} \sinh^{-1}(x/a)$$

a sets core size, s sets overall scale

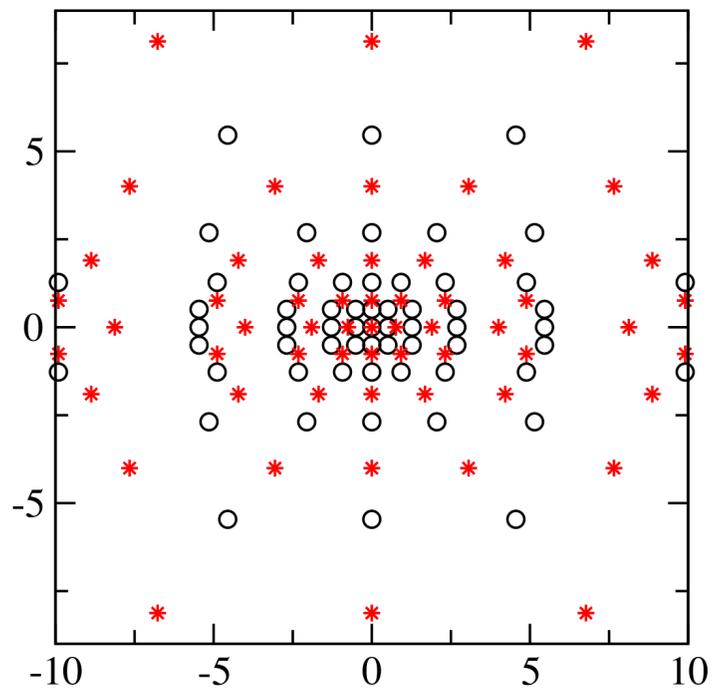


# 3D variable-grid gausslet bases

- 3D gausslets:  $G_{3D}(x, y, z) = G(x)G(y)G(z)$
- These live on a 3D distorted grid. Spacings: near nuclei,  $\sim 0.5$  Bohr, edges of atoms:  $\sim 4$  Bohr
- Add to the gausslet basis a standard Gaussian basis, which is very good near the nucleus. Orthogonalize the Gaussians to the gausslets, and you can maintain a diagonal form for the interactions

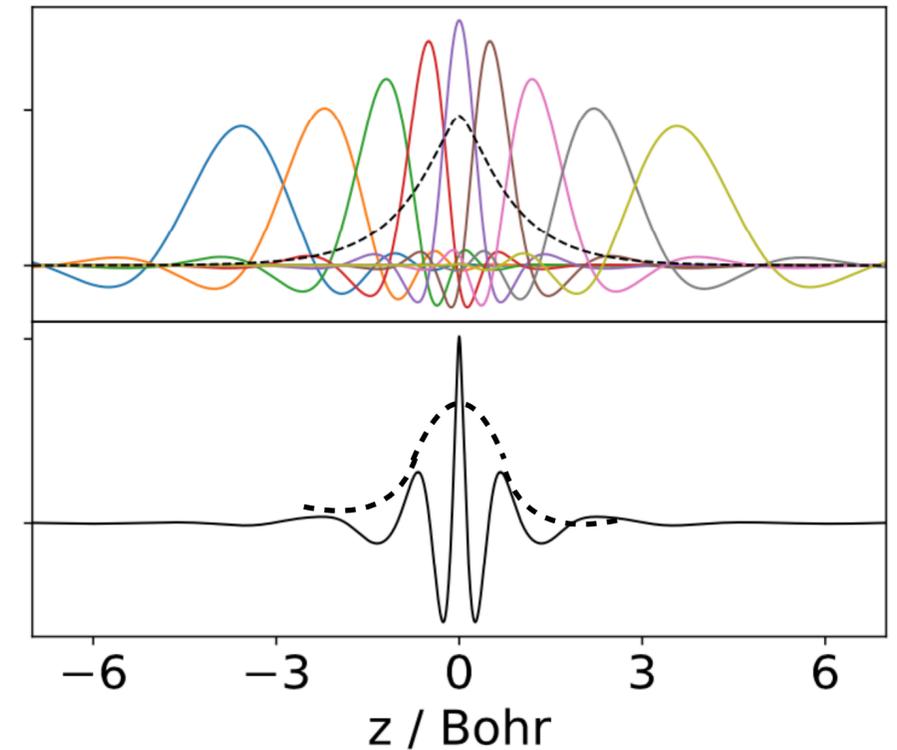


Simplest type of  
coordinate transformation



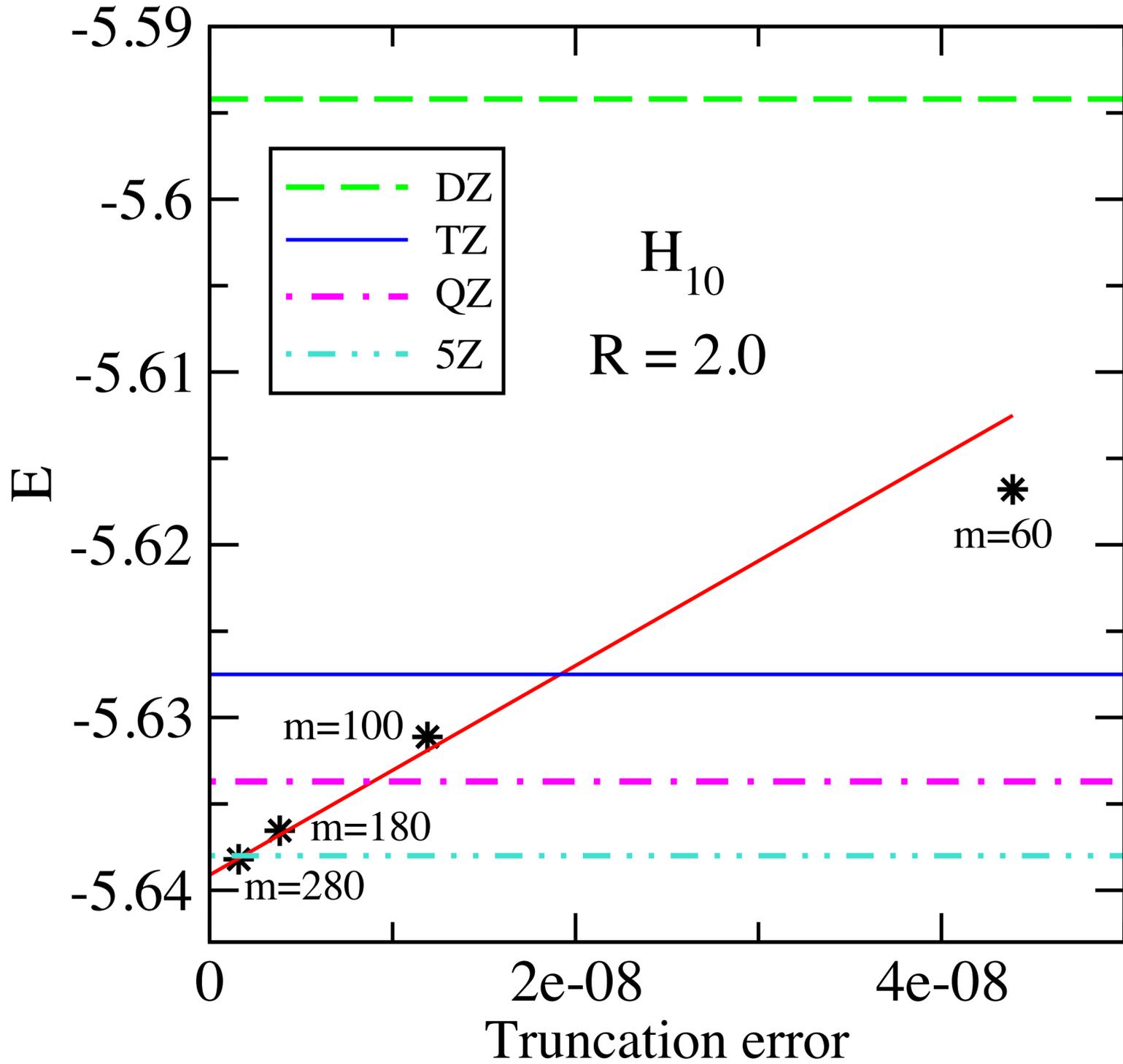
Integrals still easy

Multislicing

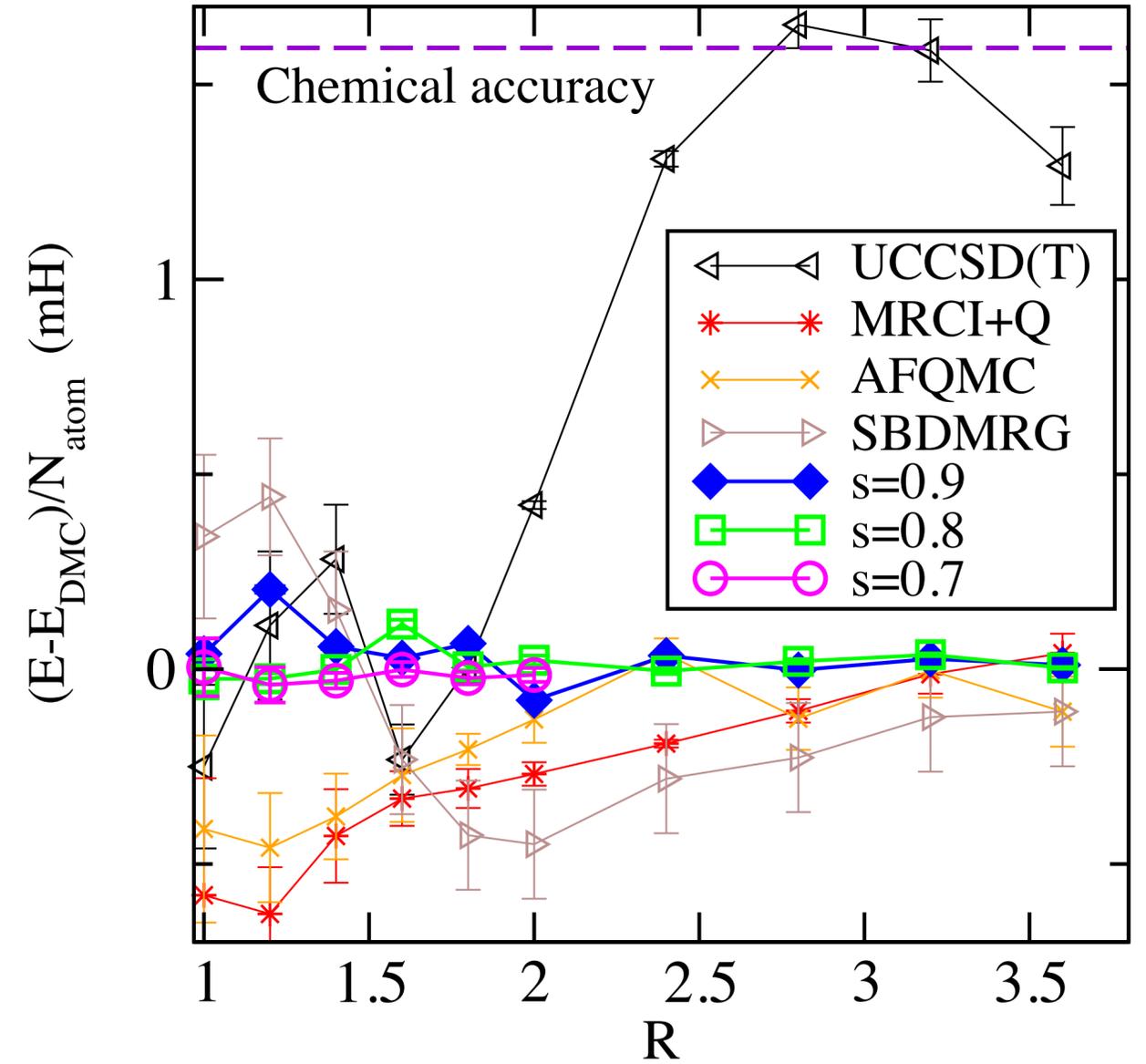


Adding in Gaussians

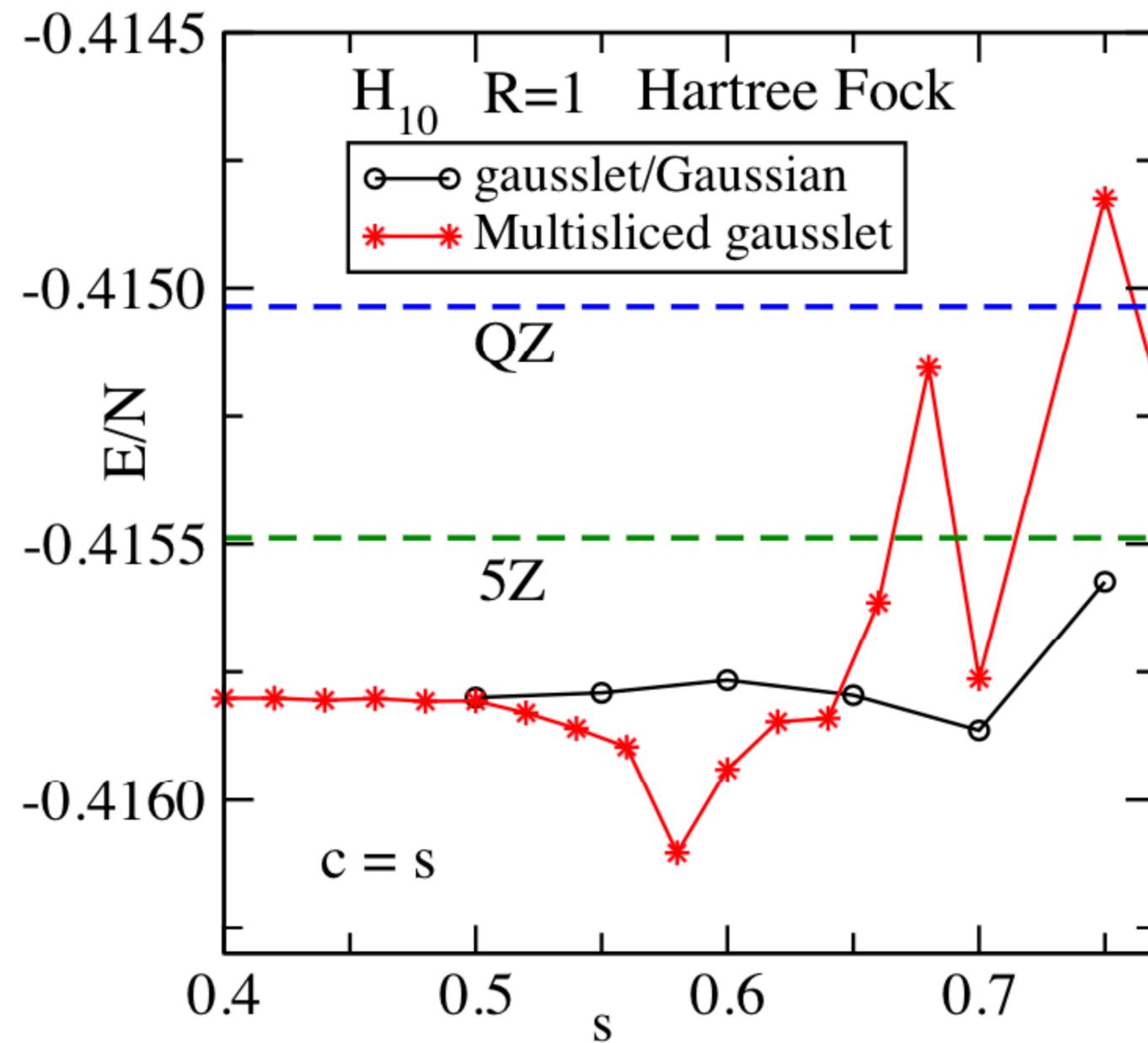
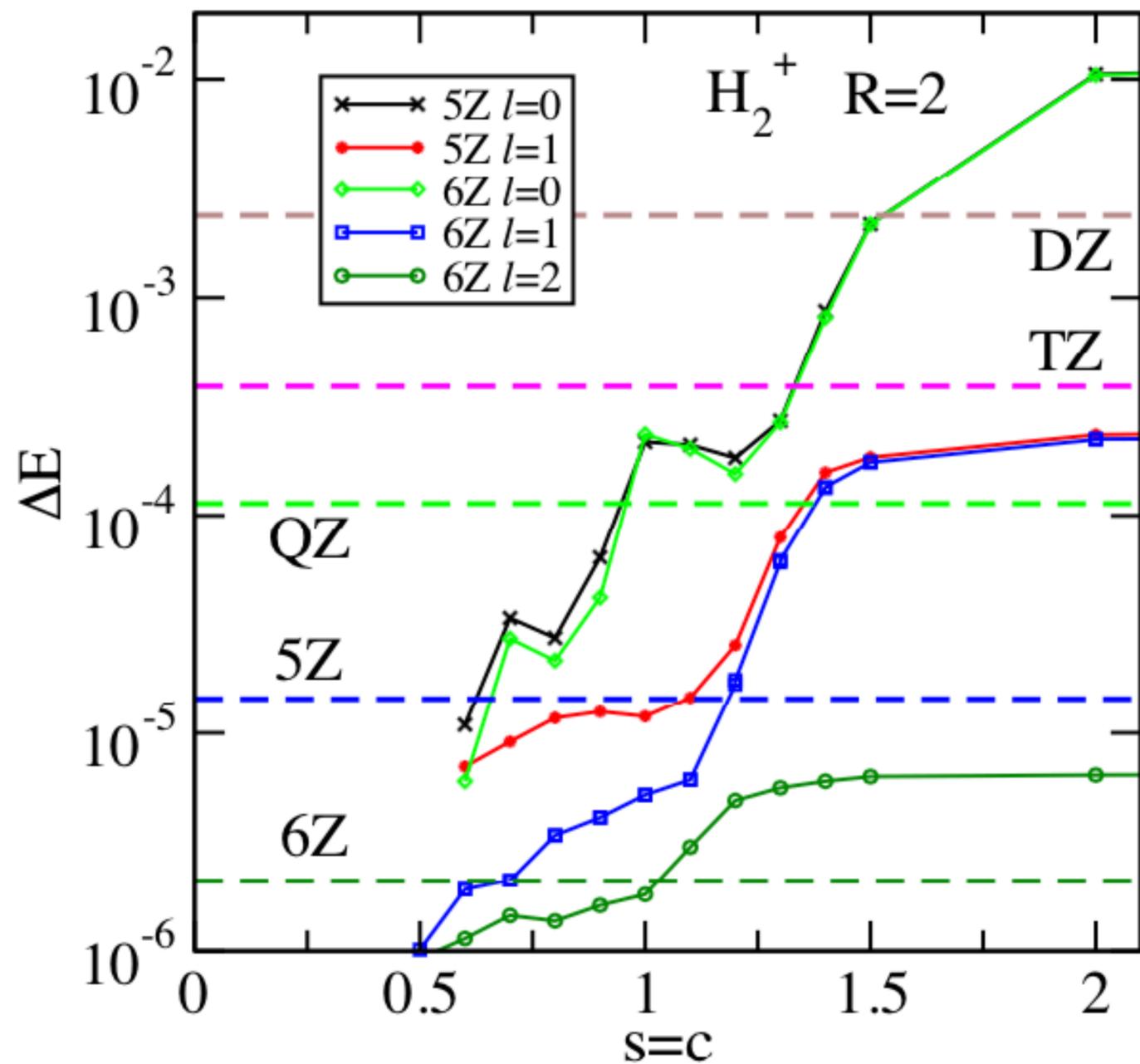
# Results: Gausslet-DMRG, H<sub>10</sub> chain



Details of DMRG calculation

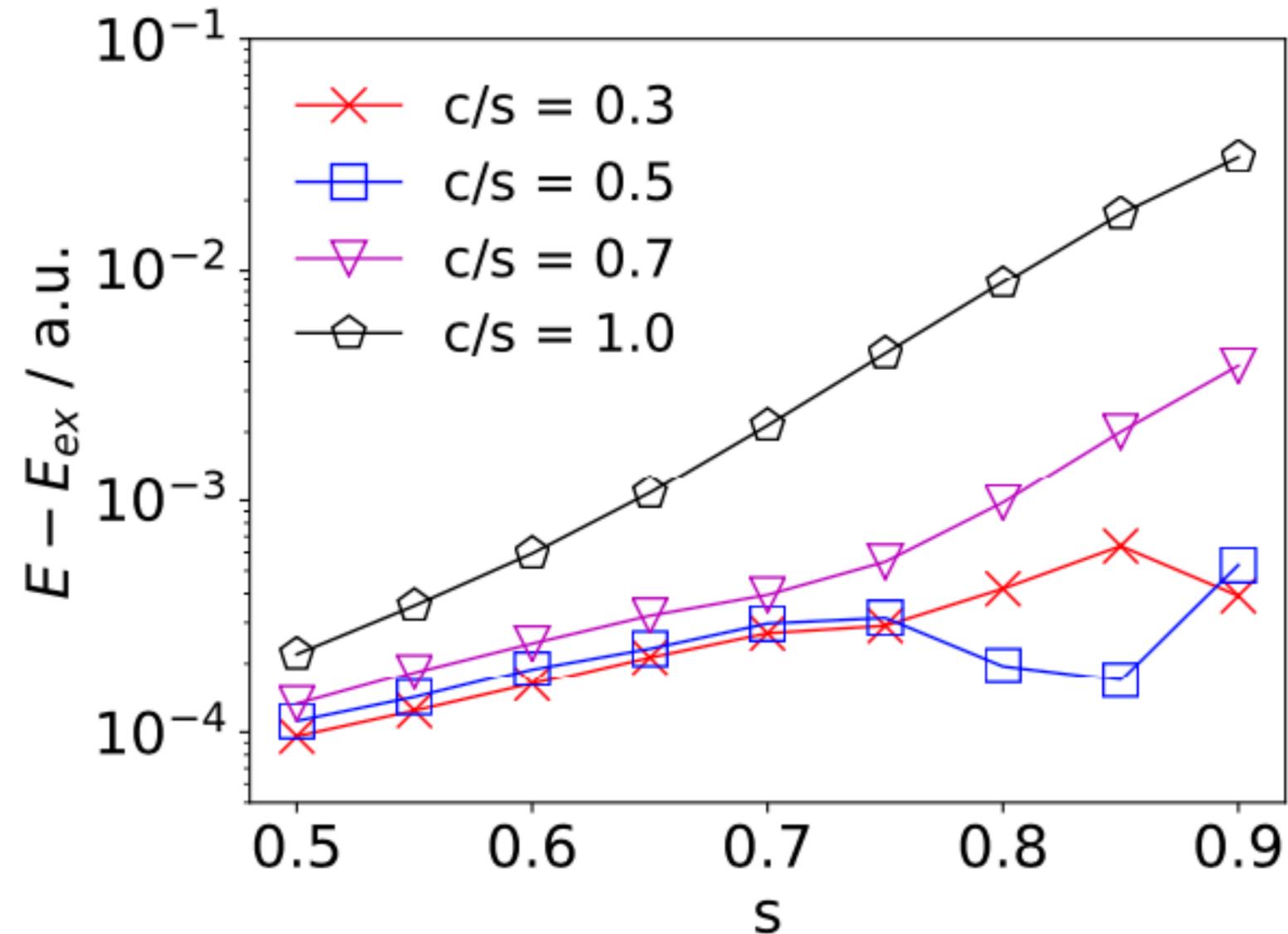


# Hybrid gausslet-Gaussian basis-DMRG (Qiu & White,...)



Errors  $\sim 10^{-5}$  a.u.

# Convergence: two electron cusp



Full CI on He atom

Easy to get below chemical accuracy,  
but also the el-el cusp slows  
convergence

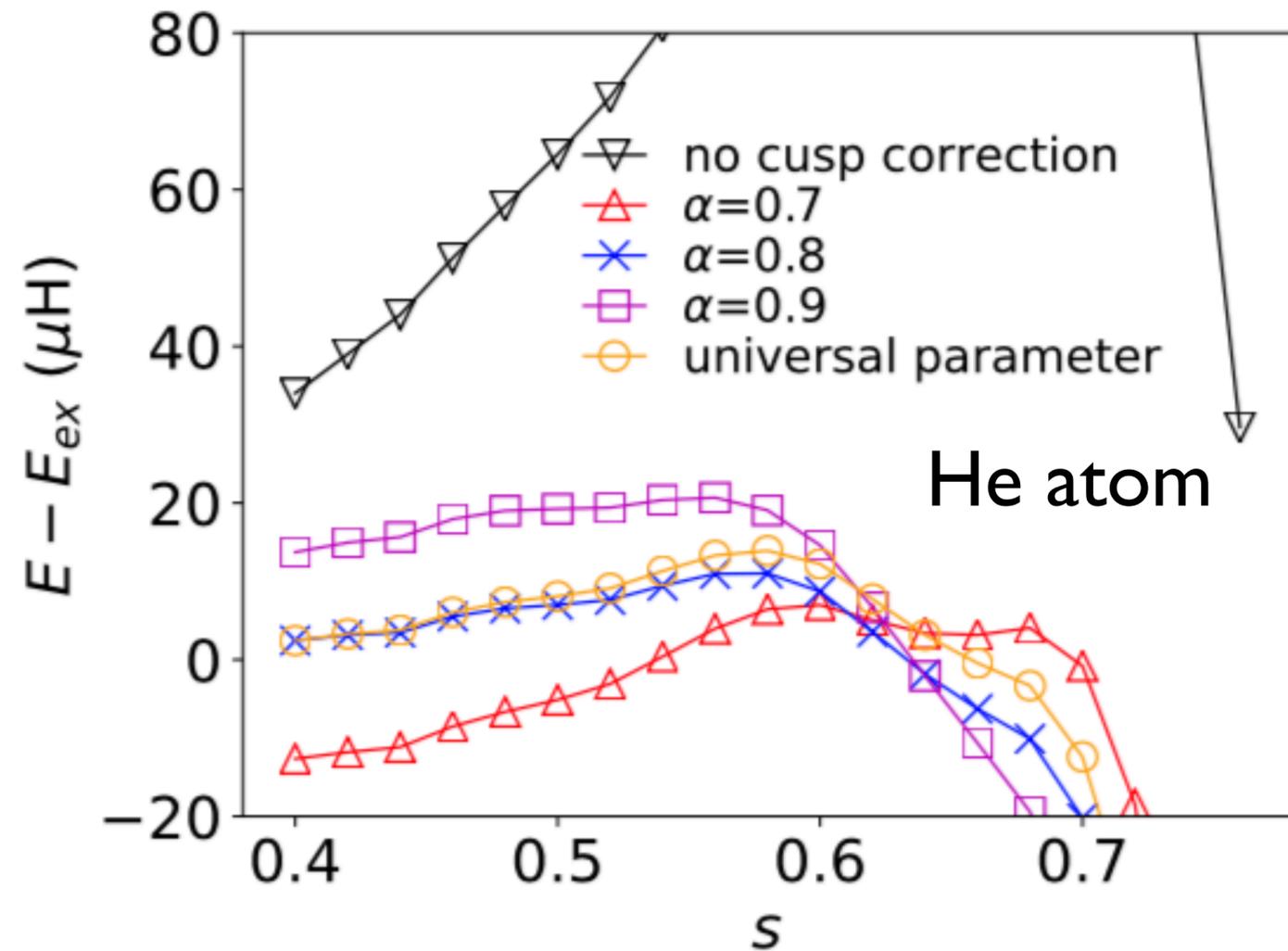
- Cusp corrections: can we correct for the el-el cusp? The local nature of gausslets makes this easier. A substantial part of the cusp appears when two electrons are on the same gausslet.

$$E_{\text{onsite}} = \sum_i V_{ii} d_i \quad \text{with} \quad d_i = \langle n_{i\uparrow} n_{i\downarrow} \rangle$$

In a complete gausslet basis,  $d_i \rightarrow 0$ . What about a  $\Delta V_{ii}$  correction? In the spirit of density functional theory (LDA), we could try to find a local correction  $\Delta V_{ii}(n_i, d_i, V_{ii})$

After studying the scaling behavior in simple cases we concluded the leading behavior was  $\Delta V_{ii}(d_i)$ , and that the asymptotic behavior looked like a power law  $\Delta E = e_0 \sum_i d_i^\alpha$  (ansatz!)

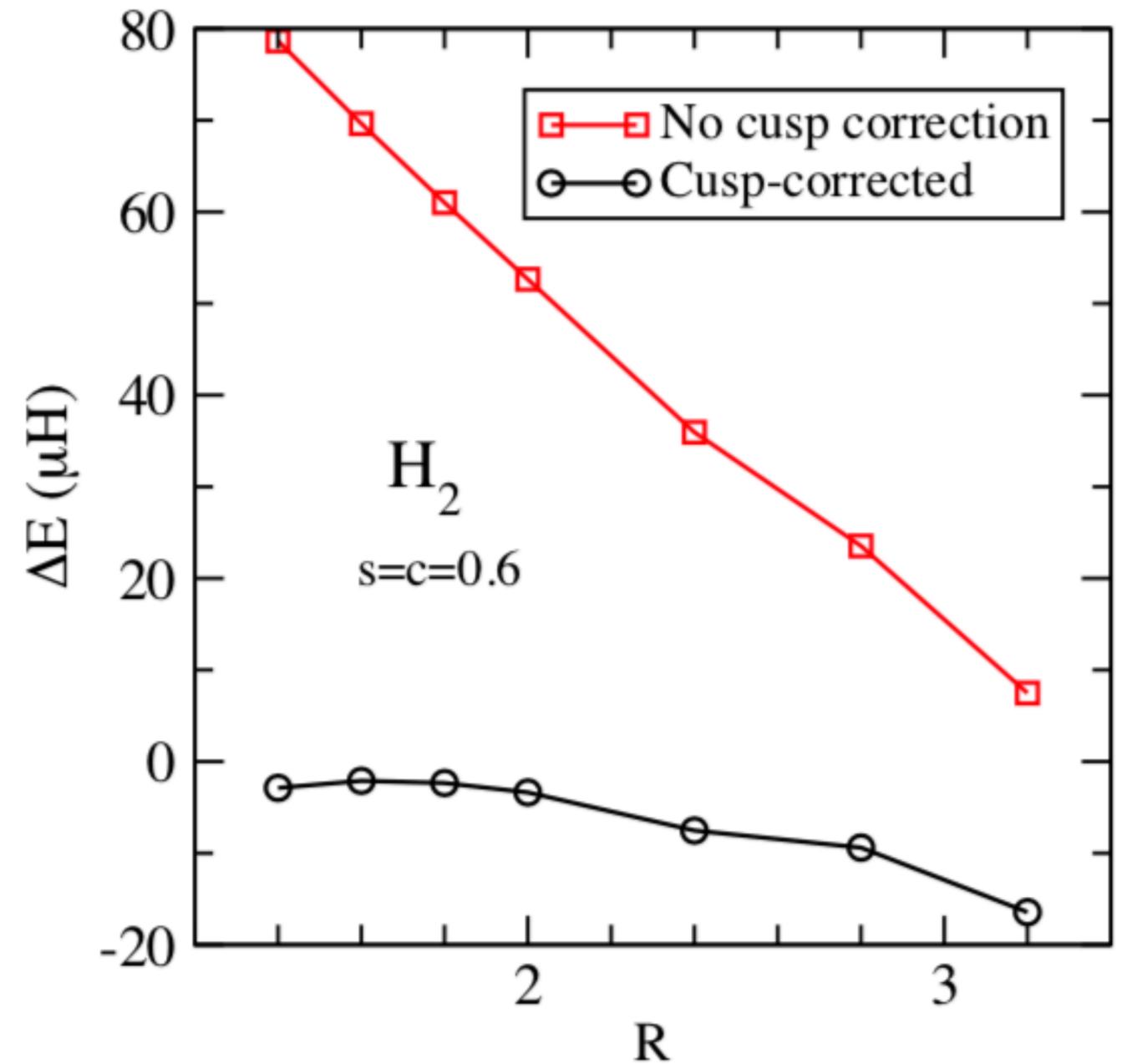
# Results: Cusp corrections



$$\Delta E = e_0 \sum_i d_i^\alpha$$

Universal values? (optimized over several two-electron systems)

$$e_0 = -0.005078 \quad \alpha = 0.79$$



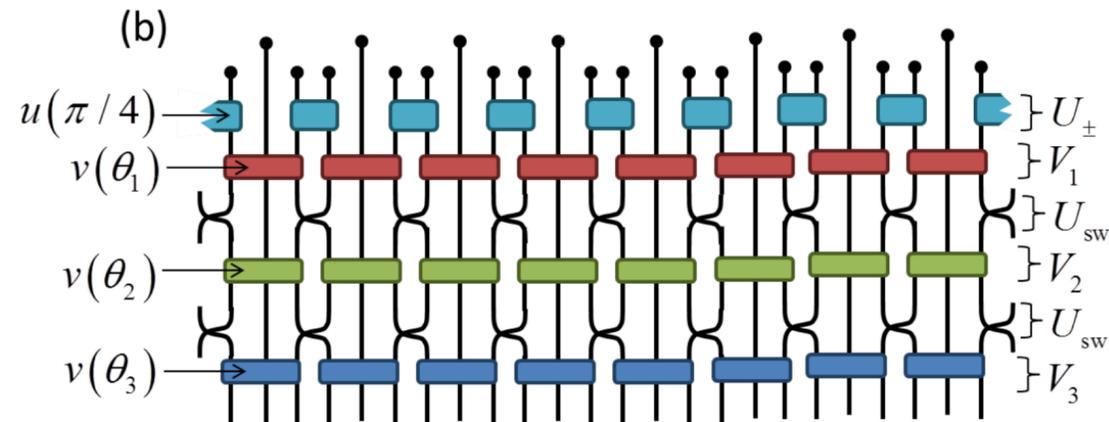
Same  $\Delta E$  parameters

# Conclusions

- How you set up the Hamiltonian has a big impact on DMRG calculations, and Gaussian basis sets are not ideal
- Sliced basis sets offer substantial advantages over Gaussians in chain systems like the H chain.
- We tested ideas about model reduction when one has fully interacting calculations available.
- Gausslets are local, smooth, orthogonal bases which give diagonal interactions
- They can be combined with ordinary Gaussians to deal with cores.
- The local nature appears to allow good el-el cusp corrections

# Gausslets—a variation on our ternary wavelets with analytic integrals

Instead of using the scale-invariant functions, apply the wavelet transform once to a array of Gaussians, with an extra “orthogonalizer” layer in between [White, J. Chem. Phys. **147**, 244102 (2017)]



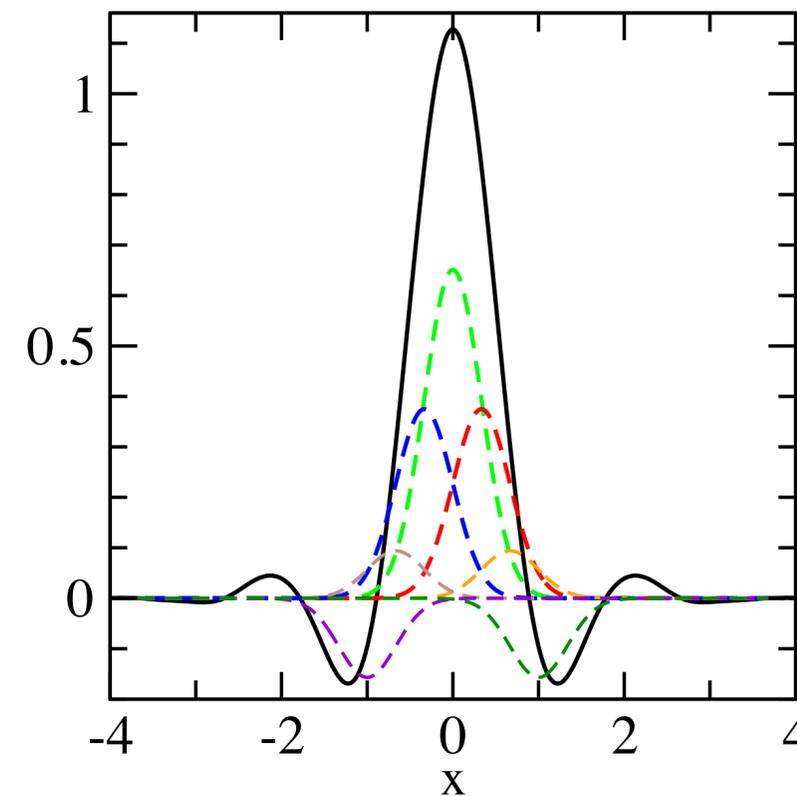
limited range approximate orthogonalizer

Uniform array of gaussians

Weeks of nonlinear optimization

$$G(x) = \sum_j b_j \exp\left[-\frac{1}{2}(3x - j)^2\right]$$

To use them, all you need is the array of  $b_j$



“Gausslet”

Put one gausslet on each grid point. Can form the bottom layer for further MERA

Exactly orthonormal

Polynomial completeness to 10th order

Integrates like a  $\delta$ -function to 20th order