Variational subspace methods

Reminder from last week

In order to compute m first excited states of H:

- Get m variational wavefunctions $|\phi_1\rangle, \dots, |\phi_m\rangle$
- They define a Slater determinant $S_- |\phi_1\rangle \otimes \dots \otimes |\phi_m\rangle$

Projector on the antisymmetric subspace $\rightarrow$ $\frac{1}{m!} \sum_{\sigma \in \mathfrak{S}_m} \text{sgn}(\sigma) P_\sigma$

$\rightarrow$ This state is unchanged upon linear combinations of $|\phi_k\rangle$

- We minimize the energy of that variational Slater determinant for the Hamiltonian $H = \sum_k H_k$

H_k: $\mathbb{1} \otimes \dots \otimes H \otimes \dots \otimes \mathbb{1}$ (H in the k-th position)

- That energy happens to be $\text{Tr}\left[G^{-1} G^{(H)}\right]$

with $G_{ij} = \langle \phi_i | \phi_j \rangle \leftrightarrow S$

$G^{(H)}_{ij} = \langle \phi_i | H | \phi_j \rangle \leftrightarrow H$

- That energy is estimated with

$$\text{Tr}\left[G^{-1} G^{(H)}\right] = \mathbb{E}_{s \sim \phi^{\otimes m}}\left[\phi^{-1}(s)\, \phi^{(H)}(s)\right]$$

with $[\phi(s)]_{ij} = \langle s_i | \phi_j \rangle$

$[\phi^{(H)}(s)]_{ij} = \langle s_i | H | \phi_j \rangle$

→ Analogous to $\langle\psi|\psi\rangle^{-1}\langle\psi|H|\psi\rangle = \mathbb{E}_{s\sim\psi}\left[\langle s|\psi\rangle^{-1}\langle s|H|\psi\rangle\right]$

– Once minimized, recover states and energies as eigenvectors and eigenvalues of $G^{-1}G^{(H)}$ | Pfaff $S^{-1}H$

→ $G^{-1}G^{(H)}\alpha^{(k)} = \mu_k \alpha^{(k)}$

→ $\begin{cases} E_k \simeq \mu_k \\ |\psi_k\rangle \simeq \sum_p \alpha_p^{(k)} |\phi_p\rangle \end{cases}$

Back to the present

(I) Subspace methods with the determinant state

(A more complete picture of NES-VMC and the determinant state)

(1) What is the determinant state, really?

Determinant state associated to the family $\{|\phi_k\rangle\}_k$

→ $|\phi_A\rangle = S_-|\phi_1\rangle\otimes\dots\otimes|\phi_m\rangle$ $\quad S_-|\phi_1\rangle|\phi_2\rangle$

We saw that a change of basis $|\psi_k\rangle = \sum_p B_{pk}|\phi_p\rangle$ does not change the determinant state

→ $|\psi_A\rangle = \det B\,|\phi_A\rangle$

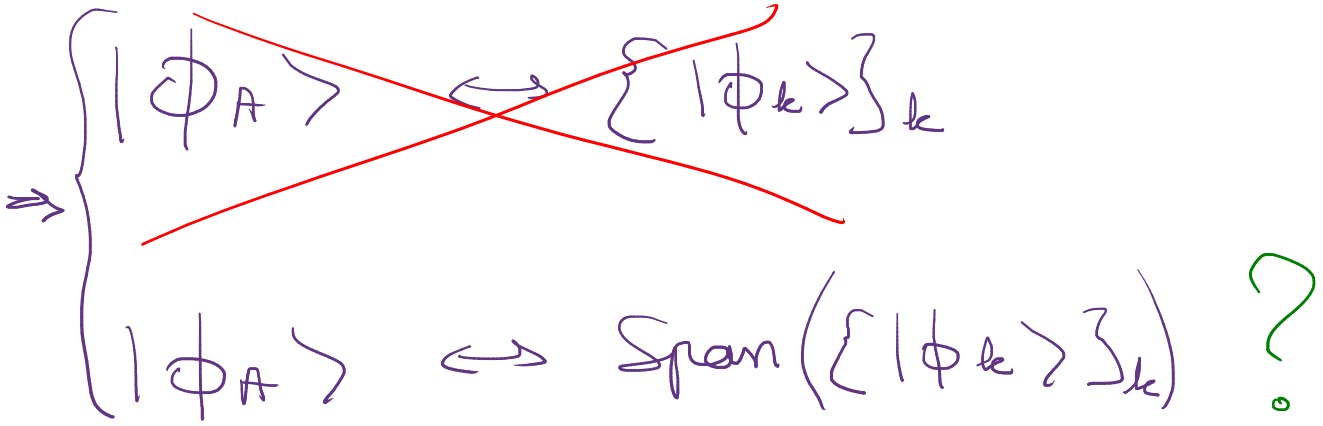
physically irrelevant

$\mathcal{H} = S_{1/2}\otimes S_{1/2}$

$|\phi_1\rangle = |+1\rangle$

$|\phi_2\rangle = |+-\rangle$

$\text{span}\{|++\rangle, |+-\rangle\} \leftrightarrow |\phi_A\rangle = S_-|\phi_1\rangle|\phi_2\rangle = |++\rangle|+-\rangle - |+-\rangle|++\rangle$



We have seen $\mathrm{Span}(\{|\phi_k\rangle\}_k) \mapsto |\phi_A\rangle$ is well-defined.

Is it injective? Does $|\phi_A\rangle$ uniquely define a subspace?

YES

Summary

$$\mathrm{Span}(\{|\phi_k\rangle\}_k) \mapsto |\phi_A\rangle$$

Is a bijective mapping from m-dimensional subspaces to Slater determinants of $\mathcal{H}^{\otimes m}$.

(Plücker embedding)

(2) Rayleigh-Ritz method and energy of a subspace

$$\frac{\langle \phi_A | H | \phi_A \rangle}{\langle \phi_A | \phi_A \rangle} = \mathrm{tr}\left(G^{-1} G^{(H)}\right)$$

with $H = \sum_k H_k$

Rayleigh-Ritz method

Ex: Lanczos algorithm

- Construct the states $|\phi_k\rangle = H^k |\phi_0\rangle$ } getting good basis states
- Orthogonormalize them $\rightarrow |\tilde{\phi}_k\rangle$
- Construct $T_{ij} = \langle \tilde{\phi}_i | H | \tilde{\phi}_j \rangle$
- $T \alpha^{(k)} = \mu_k \alpha^{(k)} \Rightarrow \begin{cases} E_k \simeq \mu_k \\ |\tilde{\Psi}_k\rangle = \sum_p \alpha_p^{(k)} |\tilde{\phi}_p\rangle \end{cases}$ } extract

Works in general for any family of orthonormal states.

If $\{|\phi_k\rangle\}_k$ is not orthonormal, we can orthonormalize with:

$$|\tilde{\phi}_p\rangle = \sum_k [G^{-1/2}]_{kp} |\phi_k\rangle \quad \text{with} \quad G_{ij} = \langle \phi_i | \phi_j \rangle$$

$$\tilde{\Phi} = \Phi G^{-1/2} \Rightarrow \tilde{\Phi}^+ \tilde{\Phi} = G^{-1/2} \Phi^+ \Phi G^{-1/2} \quad \left(\Phi_{sk} = \langle s | \phi_k \rangle\right)$$
$$= G^{-1/2} G G^{-1/2}$$
$$= \mathbb{1}$$

$$\Rightarrow \tilde{\Phi}^{+} H \tilde{\Phi} = G^{-1/2} \underbrace{\Phi^{+} H \Phi}_{G^{(H)}} G^{-1/2} \quad \text{with } G^{(H)}_{ij} = \langle \phi_i | H | \phi_j \rangle$$

In the original basis, this matrix becomes

$$G^{-1/2} G^{-1/2} G^{(H)} G^{-1/2} G^{1/2} = \boxed{G^{-1} G^{(H)}}$$

Conclusion

Rayleigh-Ritz method for non-orthonormal states

→ Diagonalize $G^{-1} G^{(H)}$ (not just $G^{(H)}$)

Guarantees of the Rayleigh-Ritz method (from min-max theorem)

$$G^{-1} G^{(H)} \alpha^{(k)} = \mu_k \alpha^{(k)}$$

→ $$E_k \leq \mu_k$$

$$\frac{\langle \psi_{\alpha^{(k)}} | H | \psi_{\alpha^{(k)}} \rangle}{\langle \psi_{\alpha^{(k)}} | \psi_{\alpha^{(k)}} \rangle} = \mu_k$$

Going back to NES-VMC:

- Optimize $\mathrm{tr}\left[G^{-1} G^{(H)}\right] = \sum_k \mu_k$ } getting good subspace
- Diagonalize $G^{-1} G^{(H)}$ } extract

→ NES-VMC is the Rayleigh-Ritz method applied to an optimized subspace

(3) Taking the determinant state beyond NES-VMC

→ NES-VMC is simply VMC applied to a subspace through the determinant state.

> Any other method can also naturally be extended to subspaces with the determinant state.

ex : Infidelity between subspaces

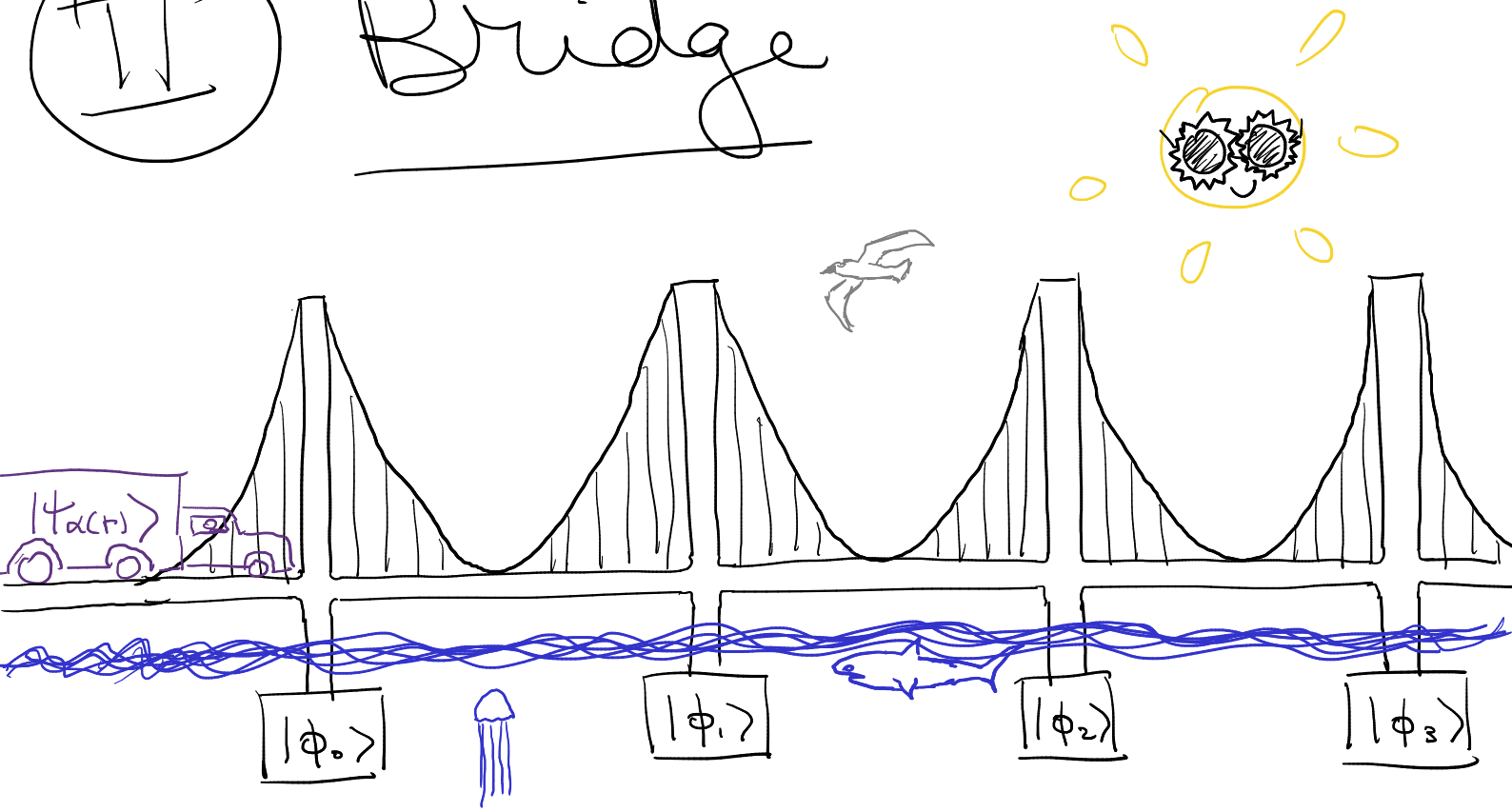
Infidelity between states : $\mathcal{F}(|\psi\rangle, |\phi\rangle) = \frac{|\langle\psi|\phi\rangle|^2}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}$

Infidelity between subspaces $U = \text{span}(\{|\psi_k\rangle\}_k)$ and $V = \text{span}(\{|\phi_k\rangle\}_k)$

$U \leftrightarrow |\psi_A\rangle$, $V \leftrightarrow |\phi_A\rangle$

→ $\mathcal{F}(U, V) = \mathcal{F}(|\psi_A\rangle, |\phi_A\rangle)$

II Bridge



TDVP / p-HVMC / etc... $\rightarrow$

$$\begin{cases} |\phi_0\rangle = |\phi_{\theta_0}\rangle \simeq |\psi(t=0)\rangle \\ |\phi_1\rangle = |\phi_{\theta_1}\rangle \simeq |\psi(t=\delta)\rangle \\ \vdots \\ |\phi_{m-1}\rangle = |\phi_{\theta_{m-1}}\rangle \simeq |\psi(t=(m-1)\delta)\rangle \end{cases}$$

Idea: Define a new variational state $|\psi_\alpha\rangle = \sum_k \alpha_k |\phi_k\rangle$ and re-run TDVP with $|\psi_\alpha\rangle$

($\theta_0, \dots, \theta_{m-1}$ are now fixed)

→ In principle, cannot be worse than the original states

→ Maybe significantly better sometimes?

TDVP in general → $S\dot{\theta} = F$ (depend on θ)

→ At every time step, recompute S and F ☹

TDVP for linear state → $\dot{\alpha} = -iG^{-1}G^{(H)}\alpha$

$G_{ij} = \langle \phi_i | \phi_j \rangle$ and $G^{(H)}_{ij} = \langle \phi_i | H | \phi_j \rangle$

do not depend on α

Only compute $G^{-1}G^{(H)}$ once ! ☺

→ Much cheaper !

Algorithm

1_ Estimate $G^{-1}G^{(H)}$

2_ Integrate $\dot{\alpha} = -iG^{-1}G^{(H)}\alpha$ exactly

(1) Estimating $G^{-1}G^{(H)}$

Estimator 1 (det)

From the determinant state perspective

$$\rightarrow \quad G^{-1}G^{(H)} = \mathbb{E}_{s\sim\phi_A}\left[\phi^{-1}(s)\,\phi^{(H)}(s)\right]$$

Estimator 2 (SOS)

single Hilbert space configuration

Sample from $p(s) = \frac{1}{\|p\|}\sum_k |\langle s|\phi_k\rangle|^2$

$$\frac{1}{\|p\|}G_{ij} = \mathbb{E}_{s\sim p}\left[\frac{\langle\phi_i|s\rangle\langle s|\phi_j\rangle}{\sum_k |\langle s|\phi_k\rangle|^2}\right]$$

(A bit like the QGT estimator)

$$\frac{1}{\|p\|}G^{(H)}_{ij} = \mathbb{E}_{s\sim p}\left[\frac{\langle\phi_i|s\rangle\langle s|H|\phi_j\rangle}{\sum_k |\langle s|\phi_k\rangle|^2}\right]$$

Comparison

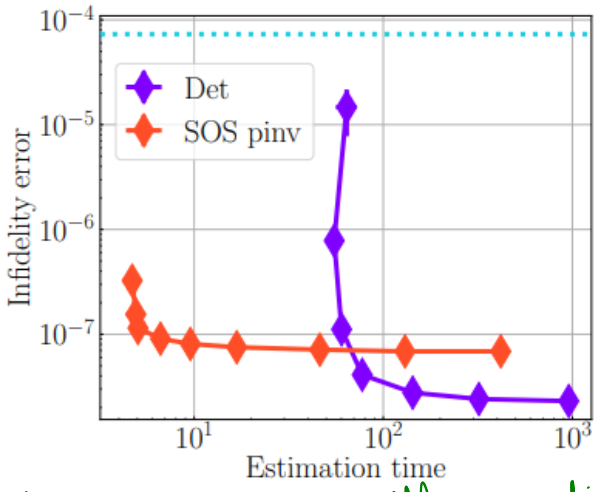


Figure 2: Comparison between the final infidelity error of Bridge performed using two estimators of the Rayleigh matrix as a function of the estimation time of the Rayleigh matrix. The two estimators are: in purple, the determinant state estimator of eq. (28), and in red, the sum of states (SOS) estimator of eqs. (39) and (40) where G is inverted using the pseudo-inverse with an appropriately chosen singular value cut-off. The dotted horizontal blue line indicates the performance of the original states $|\phi_k\rangle$. The Hamiltonian is the transverse-field Ising Hamiltonian on a 4×4 lattice with $(h, J) = (1, 0.1)$, and the basis states are 136 CNN states obtained with p-tVMC from time 0 to 1.35 with $\delta = 0.01$.

Explanation → - SOS estimator inverts an ill-conditioned matrix with Monte Carlo error
- Det only inverts matrices with numerical precision error

(2) So, does Bridge improve dynamics ?

Sometimes...

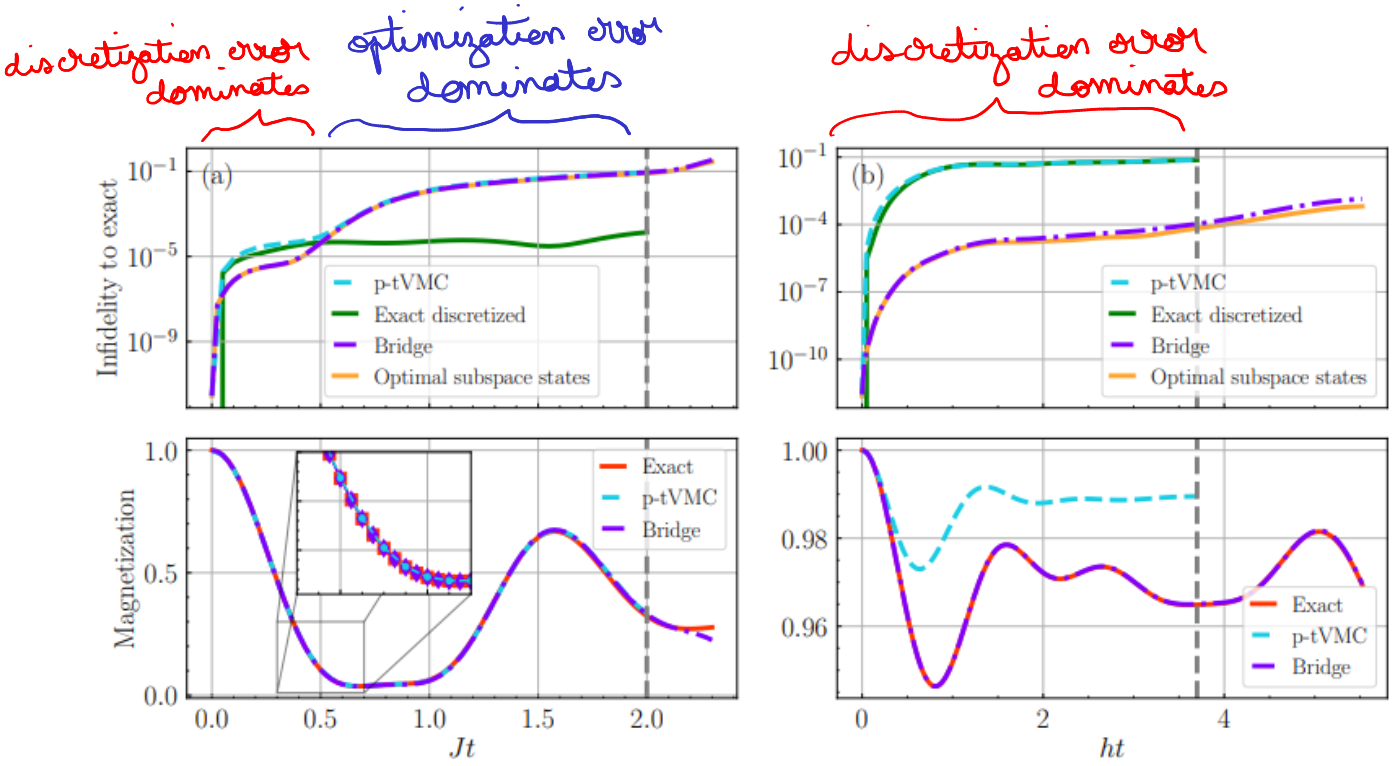


Figure 3: Infidelity to exact dynamics and magnetization on the 4×4 transverse-field Ising model with $h = 0.1h_c$ (a) and $h = 2h_c$ (b). The original states (blue) are convolutional neural networks optimized with p-tVMC using the scheme SLPE2 with time step $J\delta = 0.05$ (a) and $h\delta = 0.05$ (b). For the Bridge states (purple), the Rayleigh matrix is estimated with the determinant state estimator using 3000 samples. Exact discretized (green) refers to the dynamics resulting from the exact application of the discretized time evolution operator. The vertical grey line highlights the final time of the p-tVMC states.

Two sources of error in p-tVMC

→ Discretization error $\left(e^{-i\delta H} \simeq 1 - i\delta H + -\frac{\delta^2 H^2}{2}\right)$

→ Optimization error $\left(|\psi_0'\rangle \simeq (1 - i\delta H)|\psi_0\rangle\right)$

Bridge is helpful when discretization error dominates

Bridge summary

- Cheap (ex: 1 hour vs several days)
- Can significantly increase precision when discretization error dominates
- Can also interpolate (and extrapolate a bit)