Unitary Quantum Subspace Diagonalization – Memo 1

Ethan N. Epperly

September 19, 2022

Consider the problem of estimating the smallest eigenvalue $E = E_0$ of a large Hamiltonian matrix $\widehat{H}$ of size N. We assume that we can prepare a state $|\varphi\rangle$ with large overlap with the true ground state eigenvector $|E\rangle$, $|\langle\varphi|E\rangle| \gg 0$. Our attention will concern algorithms which obtain information from the Hamiltonian of the form

$$\alpha_j = \left\langle \varphi \middle| \mathrm{e}^{\mathrm{i}t_j\widehat{H}} \middle| \varphi \right\rangle \quad \text{for } j = 1, 2, \ldots, n$$

Such methods have two important advantages over Parrish–McMahon-type QSD algorithms:

1. They only require inner products of time-evolved states. As such, one avoids the need to make costly measurements involving the operator $\widehat{H}$.
2. These methods are robust to $\widehat{H}$ having a large numerical range. If the test vector $|\varphi\rangle$ has even a small overlap with large-eigenvalue eigenvectors of $\widehat{H}$, estimates such as $\langle\varphi|\widehat{H}|\varphi\rangle$ will suffer significant pollution from the large eigenvalues of $\widehat{H}$.

To reveal the structure of this problem, expand $|\varphi\rangle$ in an eigenbasis:

$$|\varphi\rangle = \sum_{i=0}^{N-1} \beta_i |E_i\rangle.$$

The α measurements then take the form

$$\alpha_j = \sum_{i=0}^{N-1} |\beta_i|^2 \mathrm{e}^{\mathrm{i}t_j E_i}. \tag{1}$$

Phrased in this way, our goal is to obtain the unknown quantities E_i and $|\beta_i|$ from the measurements α_j. In the typical scenario where $m \ll N$, this system will be highly undetermined and we shall only be interested in extracting a small number of E_i estimates—often only the ground state $E_0 = E$.

All but one of the methods we will consider will require the time sequence $\{t_j\}$ to be evenly spaced:

$$t_j = j\Delta t \quad \text{for } j = 1, 2, \ldots, n. \tag{2}$$

Assume this choice is made unless otherwise noted. In such cases, the measurements (and their conjugates) can be arranged into a psd Hermitian–Toeplitz matrix

$$\boldsymbol{S} = \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \overline{\alpha_1} & 1 & \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} \\ \overline{\alpha_2} & \overline{\alpha_1} & 1 & \alpha_1 & \cdots & \alpha_{n-2} \\ \overline{\alpha_3} & \overline{\alpha_2} & \overline{\alpha_1} & 1 & \cdots & \alpha_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{\alpha_n} & \overline{\alpha_{n-1}} & \overline{\alpha_{n-2}} & \overline{\alpha_{n-3}} & \cdots & 1 \end{bmatrix}.$$

This matrix has the property that $\boldsymbol{S}_{ij} = \langle\varphi|\mathrm{e}^{\mathrm{i}(t_i - t_j)}|\varphi\rangle$ using the convention $t_0 = 0$. We shall also adopt the convention that $\alpha_0 = 1$ and $\alpha_{-n} = \overline{\alpha_n}$ for $n \in \mathbb{Z}_{++}$.

1 Methods

We describe four methods for this problem. All methods require that one specify a number q of eigenvalues to estimate. These methods can be of interest both in situations where the test state $|\varphi\rangle$ is very close to an eigenvector $|\langle\varphi|E\rangle| \approx 1$ or when the overlap is nontrivial but far from unity, $|\langle\varphi|E\rangle| \approx 0.01$. The ultimate method of choice may depend on which of these scenarios we find ourselves in.

1.1 Prony's method

Prony's method begins by truncating the relation (1) for the measurements to q terms:

$$\alpha_j = \sum_{j=0}^{q-1} c_i z_i^j, \quad j = -n, -n+1, \ldots, n. \tag{3}$$

The coefficients c_i are estimates of the squared coefficients $|\beta_i|^2$ and the poles z_i are estimates for $e^{i\Delta t E_i}$. Prony's method will, unfortunately, not necessarily result in c_i which are nonnegative and z_i which lie on the unit circle, so it will help us to include the values $-n \le j \le 0$ to make the most of our data.

By the form (3) of the measurements, α_j should satisfy a degree-q recurrence

$$\alpha_j = g_1\alpha_{j-1} + g_2\alpha_{j-2} + \cdots + g_q\alpha_{j-q}.$$

Rearranging this into a linear equation for the g's we obtain

$$\begin{bmatrix} \alpha_{-n+q} \\ \alpha_{-n+q+1} \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \alpha_{-n+q-1} & \alpha_{-n+q-2} & \cdots & \alpha_{-n} \\ \alpha_{-n+q} & \alpha_{-n+q-1} & \cdots & \alpha_{-n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n-1} & \alpha_{n-2} & \cdots & \alpha_{n-q} \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_q \end{bmatrix}.$$

If $n > q$, this system of equations has at least as many equations as unknowns and is generically solvable in the least squares sense.

One can then solve for the poles z_j as the solutions to the characteristic equation

$$z^q - g_1 z^{q-1} - g_2 z^{q-2} - \cdots - g_q = 0. \tag{4}$$

The coefficients $c_0, \ldots, c_{q-1}$ can be found by solving, in the least squares sense, (3) as a linear system of equations for the coefficients $c_0, \ldots, c_{q-1}$. The solution of the characteristic equation (4) and the (Vandermonde-structured) least squares problem (3) are potentially highly ill-conditioned.

1.2 Generalized eigenvalue problem

A potentially more numerically stable approach would be to formulate the problem as a generalized eigenvalue problem. View the vectors $e^{it_j\hat{H}}|\varphi\rangle$ for $j = 0, 1, 2, \ldots, n-1$ as spanning a trial subspace $\mathcal{V}$. The Rayleigh–Ritz estimtes of the eigenvalues of the time evolution operator $e^{i\Delta t\hat{H}}$ are then given by the solution to the generalized eigenvalue problem

$$\boldsymbol{S}(1:n, 2:n+1)\boldsymbol{c} = z\boldsymbol{S}(1:n, 1:n)\boldsymbol{c}.$$

The Ritz values z should be interpreted as approximations to $e^{i\Delta t E_i}$. This is precisely the QSD variant proposed by Klymko et al. [KMC+21]. We expect that this eigenvalue problem will need to be solved using thresholding to reliably obtain accurate estimates.

1.3 Vandermonde decomposition

The problem of estimating the frequencies E_i in a mixture of the form (1) can be viewed as a matrix decomposition problem. Indeed, observe that the $\boldsymbol{S}$ matrix can be factored as

$$\boldsymbol{S} = \boldsymbol{V}_{\text{ex}} \boldsymbol{D}_{\text{ex}} \boldsymbol{V}_{\text{ex}}^*$$

where $\boldsymbol{D}_{\text{ex}} = \text{diag}(|\beta_i|^2)$ and $\boldsymbol{V}_{\text{ex}}$ is a $(n+1) \times N$ matrix with ji entry $\mathrm{e}^{\mathrm{i}(-(j-1)\Delta t)E_i}$. One approach to identifying estimates for the E_i is to compute a low-rank approximation to $\boldsymbol{S}$ of this Vandermonde product form

$$\boldsymbol{S} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^*$$

where $\boldsymbol{D}$ is a nonnegative diagonal matrix and $\boldsymbol{V}$ i a $(n+1) \times q$ Vandermonde matrix associated with poles on the unit circle. Estimates for $\mathrm{e}^{\mathrm{i}\Delta E_i}$ are just the conjugates of the second row of $\boldsymbol{V}$.

Here is a numerical procedure for obtaining such a low-rank approximation, adapted from the proof of [YX18, Thm. 1]. We begin by describing the procedure in which case $\boldsymbol{S}$ is exactly of rank q. Let

$$\boldsymbol{S} = \boldsymbol{W}\boldsymbol{W}^* \tag{5}$$

be a rank decomposition of $\boldsymbol{S}$ and let $\boldsymbol{W}_\uparrow$ and $\boldsymbol{W}_\uparrow$ denote $\boldsymbol{W}$ with its last and first rows removed, repectively. In view of the Toeplitz structure of $\boldsymbol{S}$, we have

$$\boldsymbol{S}(1:n, 1:n) = \boldsymbol{W}_\uparrow \boldsymbol{W}_\uparrow^* = \boldsymbol{W}_\downarrow \boldsymbol{W}_\downarrow^*.$$

Thus, there exists a unitary matrix $\boldsymbol{U}$ such that $\boldsymbol{W}_\downarrow = \boldsymbol{W}_\uparrow \boldsymbol{U}$. Letting $\boldsymbol{w}_j^*$ denote the jth row of $\boldsymbol{W}$, the relation $\boldsymbol{W}_\downarrow = \boldsymbol{W}_\uparrow \boldsymbol{U}$ is equivalent to the recurrence $\boldsymbol{w}_{j+1}^* = \boldsymbol{w}_j^* \boldsymbol{U}$. Thus, $\boldsymbol{w}_j^* = \boldsymbol{w}_1^* \boldsymbol{U}^{j-1}$. Diagonalize the unitary matrix $\boldsymbol{U} = \widetilde{\boldsymbol{U}}\,\text{diag}(z_1, \ldots, z_q)\widetilde{\boldsymbol{U}}^*$ where $\widetilde{\boldsymbol{U}}$ has columns $\widetilde{\boldsymbol{u}}_1, \ldots, \widetilde{\boldsymbol{u}}_q$. Then $\boldsymbol{S} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^*$ where $\boldsymbol{D}$ is the diagonal matrix with entries $|\boldsymbol{w}_1^* \widetilde{\boldsymbol{u}}_k|^2$ and $\boldsymbol{V}$ is a Vandermonde matrix with jk entry z_k^{j-1}.

To apply this procedure when $\boldsymbol{S}$ is not exactly rank q, we replace the rank decomposition (5) with an optimal rank-q approximation. For $\boldsymbol{U}$ we use the solution to the unitary procrustes problem

$$\boldsymbol{U} = \operatorname*{argmin}_{\boldsymbol{V} \text{ unitary}} \left\| \boldsymbol{W}_\downarrow - \boldsymbol{W}_\uparrow \boldsymbol{V} \right\|_{\text{F}},$$

which is given by

$$\boldsymbol{U} = \boldsymbol{U}'(\boldsymbol{V}')^* \quad \text{where } \boldsymbol{W}_\uparrow^* \boldsymbol{W}_\downarrow = \boldsymbol{U}'\boldsymbol{\Sigma}'(\boldsymbol{V}')^* \text{ is an SVD.}$$

With these two modifications, the procedure works exactly the same as when $\boldsymbol{S}$ was exactly rank-q.

1.4 Nonlinear least squares

For our last algorithm, we lift the restriction (2) that the time grid is uniform. In this approach, we attempt to learn the parameters in a sum of exponentials formula directly:

$$\alpha_j = \sum_{i=0}^{q-1} c_i \mathrm{e}^{\mathrm{i}t_j \widehat{E}_i}$$

as the solutions to a nonlinear least squares problem

$$(c_0, \ldots, c_{q-1}, \widehat{E}_0, \widehat{E}_1, \ldots, \widehat{E}_{q-1}) = \operatorname{argmin} \sum_{j=1}^{n} \left| \alpha_j - \sum_{i=0}^{q-1} c_i \mathrm{e}^{\mathrm{i}t_j \widehat{E}_i} \right|^2.$$

2 Numerical experiments

I have compared the first three algorithms on the Hubbard model example we used in the QSD paper ($L = 10$, $U = 8$, $n = 79$). I have found the nonlinear least squares algorithm to be somewhat finnicky, often getting stuck in local minima or taking a while to converge. For this reason, I do not include the nonlinear least squares algorithm in this comparison. As the error metric, I used the minimum absolute difference between the true ground state energy and *any* energy estimate produced by the algorithm. Robustly *identifying* the ground state energy may remain a problem, but if the error is small, a good approximation to the ground state was among the energy estimates produced by the algorithm. For all algorithms, I use *no thresholding or truncation*, setting $q := n$.

The average error over ten trials for the three algorithms at different noise levels is shown in Figure 1. The performance of the methods is extremely similar and all methods have errors roughly proportional to the noise level even in the limit of large amounts of error.

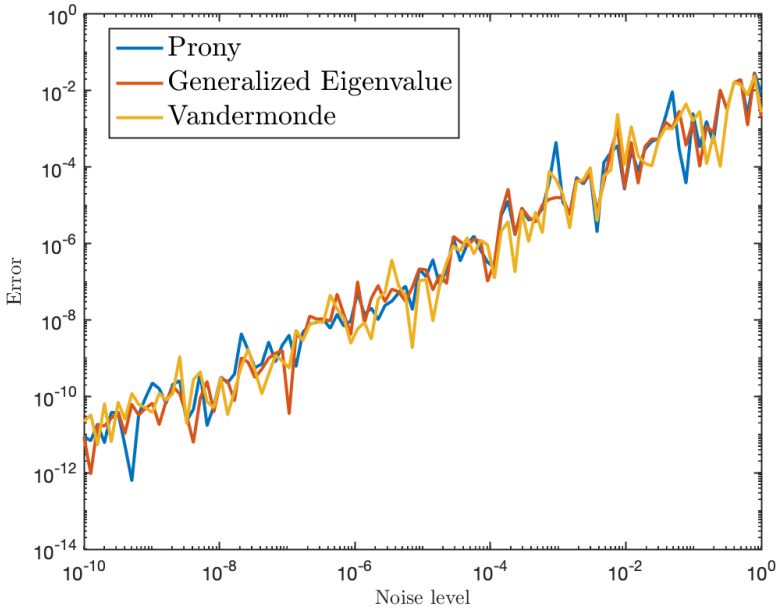

Figure 1: Comparison of different algorithms to recover the ground state energy from measuremenets α_j

To identify the correct energy estimate as the true ground state might end up becoming the true challenge for this kind of method. There are two challenges for this task:

- Spurious energies introduced by the noise.
- Genuine high energies that have been "wrapped around" to appear as small energies by the function $e^{i(\Delta t)E}$.

Perhaps these methods will fair differently at addressing these two possible failure modes.

References

[KMC+21] Katherine Klymko, Carlos Mejuto-Zaera, Stephen J. Cotton, Filip Wudarski, Miroslav Urbanek, Diptarka Hait, Martin Head-Gordon, K. Birgitta Whaley, Jonathan Moussa,

Nathan Wiebe, Wibe A. de Jong, and Norm M. Tubman. Real time evolution for ultracompact Hamiltonian eigenstates on quantum hardware. *arXiv:2103.08563 [cond-mat, physics:quant-ph]*, April 2021. 2

[YX18] Zai Yang and Lihua Xie. Frequency-selective Vandermonde decomposition of Toeplitz matrices with applications. *Signal Processing*, 142:157–167, January 2018. 3