

SUPPLEMENTARY MATERIALS: A THEORY OF QUANTUM SUBSPACE DIAGONALIZATION*

ETHAN N. EPPERLY[†], LIN LIN[‡], AND YUJI NAKATSUKASA[§]

In this supplement, we provide additional proofs and numerical experiments to support the claims made in the main text.

SM1. Proof of Theorem 4.1. For reference, we provide a complete proof of Theorem 4.1.

Proof of Theorem 4.1. Let Π and $\tilde{\Pi}$ be the spectral projectors onto the dominant m -dimensional invariant subspaces of \mathbf{A} and $\mathbf{A} + \Delta$ respectively. First, we bound

$$\begin{aligned} \|[\mathbf{A} + \Delta]_m - [\mathbf{A}]_m\|_{\text{QUI}} &= \|\tilde{\Pi}(\mathbf{A} + \Delta)\tilde{\Pi} - \Pi\mathbf{A}\Pi\|_{\text{QUI}} \\ &\leq \|\tilde{\Pi}\Delta\tilde{\Pi}\|_{\text{QUI}} + \|\tilde{\Pi}\mathbf{A}\tilde{\Pi} - \Pi\mathbf{A}\Pi\|_{\text{QUI}}. \end{aligned}$$

For the first term, we bound $\|\tilde{\Pi}\Delta\tilde{\Pi}\|_{\text{QUI}} \leq \|\Delta\|_{\text{QUI}}$ using the fact $\|\cdot\|_{\text{QUI}}$ is *symmetric* in the sense that $\|\mathbf{B}_1\mathbf{B}_2\mathbf{B}_3\|_{\text{QUI}} \leq \|\mathbf{B}_1\|_{\text{QUI}}\|\mathbf{B}_2\|_{\text{QUI}}\|\mathbf{B}_3\|_{\text{QUI}}$ [SM3, Thm. 3.9]. Every quadratic unitarily invariant norm is bounded by the Frobenius norm, which follows from the definition of quadratic unitarily invariant norm and the fact that the nuclear norm bounds every unitarily invariant norm [SM1, Eq. (IV.38)]. Thus, the second term is bounded as

$$\|\tilde{\Pi}\mathbf{A}\tilde{\Pi} - \Pi\mathbf{A}\Pi\|_{\text{QUI}} \leq \|\tilde{\Pi}\mathbf{A}\tilde{\Pi} - \Pi\mathbf{A}\Pi\|_{\text{F}},$$

which is then bounded by Theorem 2.6 with $\epsilon = \lambda_{m+1} + \|\Delta\|$ and $\rho = (\lambda_m - \lambda_{m+1} - \|\Delta\|)/(\lambda_{m+1} - \|\Delta\|)$. \square

SM2. The Failure of Heuristics. In the main text, we show how accurate recovery for the QSD algorithm usually fails, if one solves the noisy generalized eigenvalue problem (5.3) with no special treatment. A forthcoming example (Figure SM6) shows that using a fixed threshold independent of the noise level may not perform much better. An alternate strategy, which we initially believed to be more promising than thresholding, is to compute the eigenvalues (either with no thresholding at all or with a small threshold independent of the noise level) and attempt to determine real from spurious eigenvalues by means of some property of the computed eigenvector. In this section, we shall consider a couple variants of such an approach and ultimately conclude the performance of these heuristics can still be unsatisfactory when compared to thresholding.

Two natural heuristics for whether \tilde{E} is a plausible candidate for the ground state energy suggest themselves. Let $\tilde{\mathbf{c}}$ be the unit-norm eigenvector associated with a computed eigenvalue \tilde{E} . The Ritz vector $\tilde{\psi}_0 := \sum_{j=0}^{n-1} \tilde{\mathbf{c}}_j \boldsymbol{\varphi}_j$ is supposed to be close to the true ground-state eigenvector ψ_0 of \hat{H} . Our heuristics are as follows:

*Supplementary material for SIMAX MS#M145954.

<https://doi.org/10.1137/21M145954X>

[†]Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125 USA (epperly@caltech.edu).

[‡]Department of Mathematics, and Challenge Institute of Quantum Computation, University of California Berkeley, Berkeley, CA 94720 USA, and Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (linlin@math.berkeley.edu).

[§]Mathematical Institute, Oxford University, Oxford, UK (nakatsukasa@maths.ox.ac.uk).

1. **Require $h_1 := \tilde{\mathbf{c}}^* \tilde{\mathbf{S}} \tilde{\mathbf{c}}$ to be large.** The squared norm of the Ritz vector is precisely $\tilde{\mathbf{c}}^* \tilde{\mathbf{S}} \tilde{\mathbf{c}} \approx h_1$. If h_1 is small, then the norm of the Ritz vector is very small due to cancellations in the sum $\sum_{j=0}^{n-1} \tilde{\mathbf{c}}_j \boldsymbol{\varphi}_j$ and should thus be treated as suspect because of the noise. Thus, it is natural to insist on a large value of h_1 .¹
2. **Require the estimated overlap $h_2 := |\mathbf{e}_0^* \tilde{\mathbf{S}} \tilde{\mathbf{c}}| \approx |\boldsymbol{\varphi}_0^* \tilde{\boldsymbol{\psi}}_0|$ to be large.** It is important that the initial vector $\boldsymbol{\varphi}_0$ has a relatively large initial overlap $|\boldsymbol{\varphi}_0^* \boldsymbol{\psi}_0|$ with the eigenvector of interest—indeed, our analysis suggests accurate recovery of the ground-state energy requires this (see Theorem 3.1). As such, it is natural to use the overlap (or its surrogate h_1 computable from the noise-corrupted $\tilde{\mathbf{S}}$ matrix) as a measure of whether an eigenvalue is a genuine candidate for the ground-state energy. Note that by unit-norm scaling $\tilde{\mathbf{c}}$ (rather than adopting the normalization $\tilde{\mathbf{c}}^* \tilde{\mathbf{S}} \tilde{\mathbf{c}} = 1$), we are implicitly also incorporating the condition for $\tilde{\boldsymbol{\psi}}_0$ to be a stable linear combination of the basis states which motivated our interest in h_1 .

There are several ways of using a heuristic $h \in \{h_1, h_2\}$ as an algorithm for computing the ground-state eigenvalue: (a) pick E with the highest h , (b) pick the smallest E of the eigenvalues with the top k values of h , and (c) pick the smallest E with h above some thresholding h_0 (or simply the largest h if none exceeds h_0).

Unfortunately, unlike thresholding where there is a natural choice of the threshold parameter (related to the noise level η (2.1) which can usually be reliably estimated), we are unaware of any good systematic ways to pick the parameters k and h_0 for strategies (b) and (c). These heuristics thus usually require some tuning to make them accurate for a given problem instance, with the parameters needing to be readjusted when a new problem is encountered. This reduces the reliability of these heuristics, when the ground truth is unavailable to compare against. The robustness of heuristics such as (a), (b), and (c) can be improved by medians of repeated trials or by comparing the results of different heuristics against each other. However, even with such improvements, without rigorous guarantees, the validity of these heuristics remains conjectural when the genuine ground-state energy is unavailable to be validated against.

Figure SM1 shows the suggested heuristics (a), (b), and (c) with the figures of merit h_1 and h_2 with $k = 5$ and $h_0 = 10^{-2} \|\tilde{\mathbf{S}}\|$. The first subfigure, Figure SM1a, shows a relatively optimistic case for the heuristics. For low levels of noise, the eigenvalue is generally recovered with low error with the exception of a few outliers which could be ameliorated by the median trick. Figure SM1b shows the potential danger of applying these heuristics; despite working well for the Hubbard example with $n = 20$ in Figure SM1a, the heuristics fail with the same parameter choices for the Ising example with $n = 40$. For this problem, the eigenvalues are observed to be recovered accurately only with very small probability. Improvements to both plots are likely possible by more careful choice of the heuristic parameters or more complicated heuristics, but this is a point against such heuristics rather than for them: we ideally want a method which works well without tuning problem-dependent parameters.

SM3. Automatic Thresholding. As we saw in the main text, choosing a good maximum thresholding level ϵ is critical to the success of the thresholding procedure Algorithm 1.1. A useful “sanity check” is thus to solve the problem using a handful of plausible thresholding parameters to make sure the computed eigenvalues are close to

¹ h_1 is also related to the conditioning of the eigenvalue [SM3, Eq. (VI.2.2)].

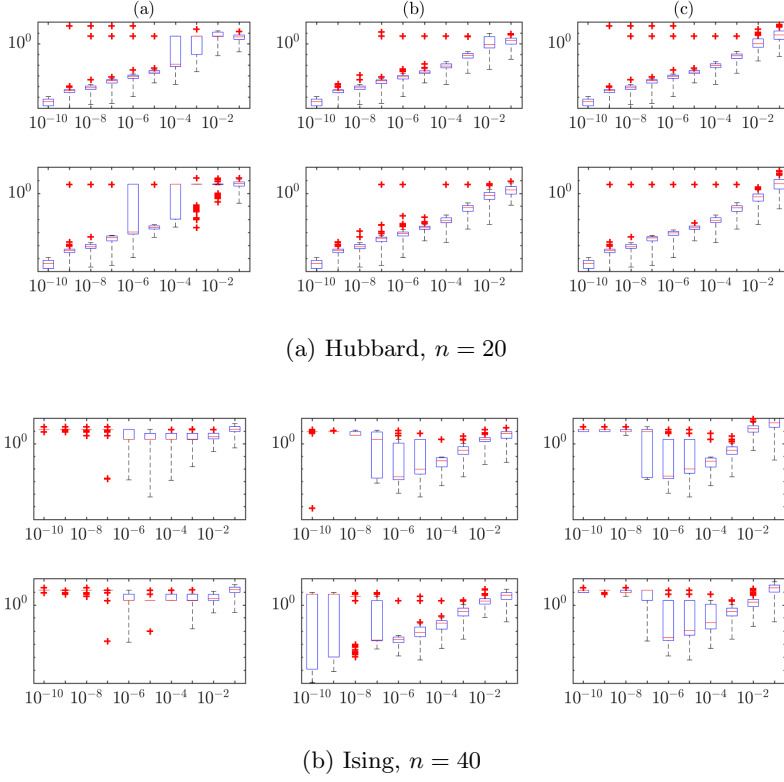


FIG. SM1. Errors (vertical axis) for eigenvalues computed from the perturbed pair $(\widetilde{\mathbf{H}}, \widetilde{\mathbf{S}})$ with heuristics (a), (b), and (c) described in the text for quality metrics h_1 (first row) and h_2 (second row). The noisy generalized eigenvalue problem (5.3) is solved using thresholding with a fixed threshold parameter $\epsilon = 10^{-12}\|\widetilde{\mathbf{S}}\|$. Shown are 100 random initializations of the noise for several random noise levels σ (horizontal axis) for the Hubbard example with $n = 20$ (top) and Ising example with $n = 40$ (bottom).

Algorithm SM3.1 Automatically tuned thresholding procedure for finding the least eigenvalue of a noise-corrupted generalized eigenvalue problem.

```

procedure AUTOTHRESHOLDING( $\mathbf{H}, \mathbf{S}, \epsilon_0, r$ )
   $E \leftarrow \text{THRESHOLDING}(\mathbf{H}, \mathbf{S}, \epsilon_0)$ 
   $\Lambda \leftarrow \{\lambda \in \text{eig}(\mathbf{S}) : \lambda < \epsilon_0\}$ 
  while  $\Lambda \neq \emptyset$  do
     $\epsilon \leftarrow \max \Lambda, \Lambda \leftarrow \Lambda \setminus \{\epsilon\}$ 
     $E' \leftarrow \text{THRESHOLDING}(\mathbf{H}, \mathbf{S}, \epsilon)$ 
    if  $|E - E'| / \min(|E|, |E'|) > r$  then
      break
    end if
     $E \leftarrow E'$ 
  end while
  return  $E$ 
end procedure

```

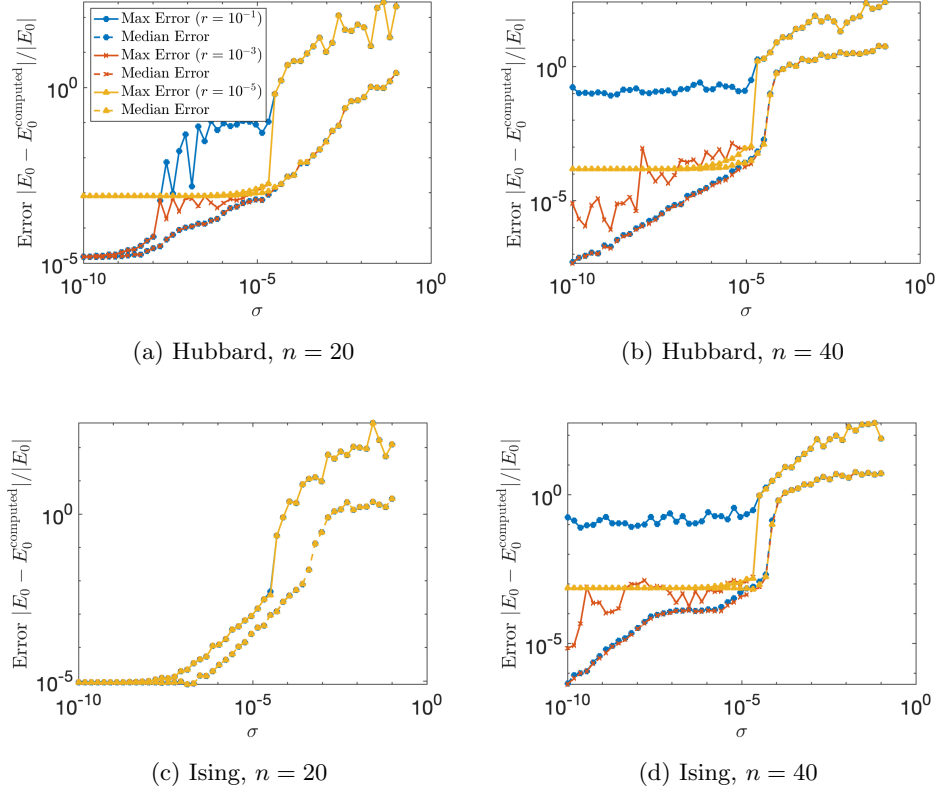


FIG. SM2. Maximum and median error over 100 initializations for eigenvalues computed from the noise-perturbed pair $(\tilde{\mathbf{H}}, \tilde{\mathbf{S}})$ using automatically tuned thresholding (Algorithm SM3.1) for three cutoffs $r \in \{10^{-1}, 10^{-3}, 10^{-5}\}$ for various values of the noise level σ for Hubbard model (top) and Ising model (bottom) with $n = 20$ (left) and 40 (right).

each other. (A variant of this strategy is proposed by Parlett for the Fix–Heiberger procedure [SM2, section 15.5].) A more ambitious strategy is to solve with a range of thresholding parameters beginning with a conservative (but not comically large) threshold parameter ϵ_0 and then tuning it down until the eigenvalue “jumps” to a presumably spurious value. The best approximation to the ground-state energy suggested by this procedure is the last value before this jump. If one wishes to automate this procedure, one needs to have a mechanistic way of deciding whether a jump has occurred: For this purpose, we shall test whether the relative difference exceeds a cutoff r . This procedure is demonstrated in Algorithm SM3.1. The success of this procedure relies on the choice of ϵ_0 not being too large as the method uses the eigenvalue recovered with parameter ϵ_0 as a baseline, large deviations from which are characterized as erroneous. Usually, one will have some good estimate of the amount of noise so picking a sensible ϵ_0 should be possible.

The performance of the automatic thresholding procedure Algorithm SM3.1 with three choices of the parameter r are shown in Figure SM2. These plots represent the worst-case situation where the noise level is completely unknown and the choice one has available for ϵ_0 is a constant multiple of $\|\tilde{\mathbf{S}}\|$. The best case scenario is shown in the $r = 10^{-3}$ lines in Figures SM2b and SM2c; in these cases, the error decays nicely as the noise does with the procedure being relatively robust (as shown

by the error over a maximum over 100 trials being similar to the median). This automatic thresholding procedure can still be somewhat delicate, with the maximum error over 100 runs being near the cutoff r for the $r = 10^{-1}$ in Figures SM2a, SM2b, and SM2c; this suggests, in the worst case, one must be willing to accept an error level on the order r due to overly aggressive automatic tuning of the thresholding parameter. However, these same plots illustrate the importance of not being too cautious either, with the maximum error being $\approx 10^{-3}$ with $r = 10^{-5}$ due to overly conservative automatic tuning of the thresholding parameter. In totality, Figure SM2 shows that the automatic thresholding procedure cannot determine a near-optimal choice for the thresholding parameter in all cases, but it can be useful in “upgrading” an overly cautious threshold parameter ϵ_0 to a better choice for ϵ , obtaining a couple more decimal digits of accuracy in the best case. As a final comment, observe that thresholding is an inherently discrete process since each eigenvalue must either be discarded or not. Therefore, even with automatically tuned thresholding, there can be plateaus in the noise-level vs accuracy curve, owing to the importance of a single eigenvalue to the overall error; this is shown in Figure SM2c.

SM4. Validation of Theorems 3.1 and 4.2. Despite its desirable theoretical implications, Theorem 3.1 may still significantly overestimate the error incurred by thresholding in practice. Consider the following examples:

- (I) $\widehat{\mathbf{H}} = \widehat{\mathbf{H}}_1$ and $\boldsymbol{\varphi}_0 = \boldsymbol{\xi}_I$,
- (II) $\widehat{\mathbf{H}} = \widehat{\mathbf{H}}_1$ and $\boldsymbol{\varphi}_0 = \boldsymbol{\xi}_{II}$,
- (III) $\widehat{\mathbf{H}} = \widehat{\mathbf{H}}_2$ and $\boldsymbol{\varphi}_0 = \boldsymbol{\xi}_{III}$,
- (IV) $\widehat{\mathbf{H}} = \widehat{\mathbf{H}}_2$ and $\boldsymbol{\varphi}_0 = \boldsymbol{\xi}_{IV}$.

where

$$\begin{aligned}\boldsymbol{\xi}_I &= \left(\sqrt{1 - 10^{-4}}, \sqrt{\frac{10^{-4}}{998}}, \sqrt{\frac{10^{-4}}{998}}, \dots, \sqrt{\frac{10^{-4}}{998}} \right) \in \mathbb{C}^{999}, \\ \boldsymbol{\xi}_{II} &= \left(\sqrt{0.5}, \sqrt{\frac{0.5}{998}}, \sqrt{\frac{0.5}{998}}, \dots, \sqrt{\frac{0.5}{998}} \right) \in \mathbb{C}^{999}, \\ \boldsymbol{\xi}_{III} &= \left(\sqrt{1 - 10^{-4} - 10^{-8}}, \sqrt{\frac{10^{-4}}{998}}, \sqrt{\frac{10^{-4}}{998}}, \dots, \sqrt{\frac{10^{-4}}{998}}, 10^{-4} \right) \in \mathbb{C}^{1000}, \\ \boldsymbol{\xi}_{IV} &= \left(1, \frac{0.01}{2}, \frac{0.01}{3}, \dots, \frac{0.01}{1000} \right) / \left\| \left(1, \frac{0.01}{2}, \frac{0.01}{3}, \dots, \frac{0.01}{1000} \right) \right\| \in \mathbb{C}^{1000},\end{aligned}$$

and

$$\begin{aligned}\widehat{\mathbf{H}}_1 &= \text{diag} \left(1, 2 + 0 \cdot \frac{0.1}{997}, 2 + 1 \cdot \frac{0.1}{997}, \dots, 2 + 997 \cdot \frac{0.1}{997} \right) \in \mathbb{C}^{999 \times 999}, \\ \widehat{\mathbf{H}}_2 &= \text{diag} \left(1, 2 + 0 \cdot \frac{0.1}{997}, 2 + 1 \cdot \frac{0.1}{997}, \dots, 2 + 997 \cdot \frac{0.1}{997}, 1000 \right) \in \mathbb{C}^{1000 \times 1000}.\end{aligned}$$

These examples are artificial: They are engineered to have a large spectral gap ΔE_1 but a small spectral range ΔE_{999} . Even with these artificial examples, Theorem 3.1 (as well as Theorems 3.1 and 4.2 together) still overestimates the error by several orders of magnitude. See Figure SM3.

We consider the bound Theorem 4.2 by itself in Figure SM4. Shown is the error $\widehat{E}_0 - E_0$ between the recovered least eigenvalue \widehat{E}_0 and the true least eigenvalue E_0 for different choices of threshold parameter ϵ . For this example, we used

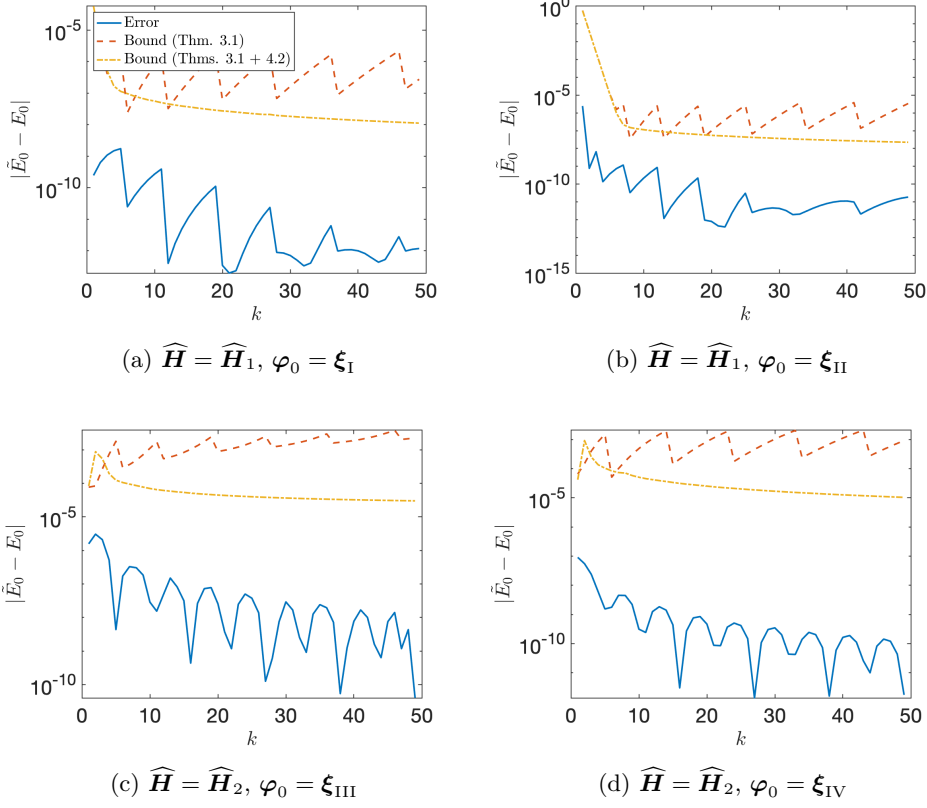


FIG. SM3. Error for QSD with the time sequence from the hypotheses of Theorem 3.1 with threshold parameter $\epsilon = 10^{-6}$ for various values k and four different sets of input data.

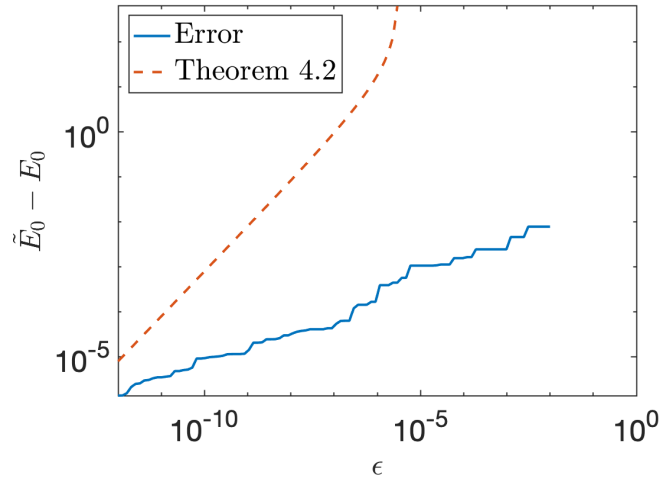


FIG. SM4. Error due to thresholding and error bound from Theorem 4.2 for least eigenvalue computed from a synthetically generated pair (\mathbf{H}, \mathbf{S}) for various threshold parameters ϵ .

$\mathbf{H} = \mathbf{K}^* \text{diag}(1, 2, \dots, 100) \mathbf{K}$ and $\mathbf{S} = \mathbf{K}^* \mathbf{K}$, where \mathbf{K} is a product of diagonal matrix with (j, j) th entry j^{-2} and a “randsvd” matrix from MATLAB’s gallery with approximate condition number 10^3 . This was chosen to give a fairly ill-conditioned “ \mathbf{S} ” matrix ($\kappa(\mathbf{S}) \approx 10^{12}$) in which the \mathbf{S} -normalized ground state eigenvector of fairly small norm ($\|\mathbf{c}_0\| \approx 10^2$). We find the bound from Theorem 4.2 becomes increasingly conservative as ϵ increases, diverging to $+\infty$ at $\epsilon \approx 10^{-6}$ which the true error $\tilde{E}_0 - E_0$ remains bounded $\leq 10^{-2}$ for $\epsilon \leq 10^{-2}$.

SM5. Evidence for Tightness of Theorem 2.5. First, we present a synthetically generated numerical example which suggests that the η/ϵ^α behavior in Theorem 2.5 is necessary, at least without further assumptions. As our example, we set $\mathbf{A} = (\mathbf{G} + \mathbf{G}^*)/2$ to be the Hermitian part of a 5×5 real standard Gaussian matrix \mathbf{G} and pick $\mathbf{S} = \text{diag}(1, 0.1, 3 \times 10^{-10}, 2 \times 10^{-10}, 10^{-10})$ and $\mathbf{H} = \mathbf{S}^{1/2} \mathbf{A} \mathbf{S}^{1/2}$. By construction, this example obeys the geometric mean bound (2.12) with $\alpha = 1/2$ and $\mu = 0.5 \lambda_{\max}(\mathbf{G} + \mathbf{G}^*)$ which is $\lesssim 10$ with high probability. We choose a threshold level of $\epsilon = 1.5 \times 10^{-10}$, so that the thresholded problem has dimension four. As perturbation, we take $\Delta_{\mathbf{S}} = 10^{-12} \cdot (\mathbf{\Gamma} + \mathbf{\Gamma}^*)/2$ (for a 5×5 real standard Gaussian matrix $\mathbf{\Gamma}$).

Let Π and $\tilde{\Pi}$ denote the spectral projectors onto the eigenvectors $> \epsilon$ for \mathbf{S} and $\tilde{\mathbf{S}} = \mathbf{S} + \Delta_{\mathbf{S}}$ respectively. For one random initialization of the Gaussian test matrices (which we find is broadly representative of repeat trials), we computed

$$(SM5.1) \quad \|\tilde{\Pi} \mathbf{H} \tilde{\Pi} - \Pi \mathbf{H} \Pi\| = 6.3 \times 10^{-8} \approx 10^{-7} \approx \|\Delta_{\mathbf{S}}\|/\epsilon^{1/2}.$$

Were the $\epsilon^{-\alpha}$ dependence in Theorem 2.5 unnecessary, we would expect that the projection error $\|\tilde{\Pi} \mathbf{H} \tilde{\Pi} - \Pi \mathbf{H} \Pi\|$ would be bounded by $\mu(1 + \rho^{-1})5^3 \|\Delta_{\mathbf{S}}\| \approx 10^{-9}$. We take this as evidence that the $\epsilon^{-\alpha}$ factor in the bound in Theorem 2.5 is necessary, at least without additional assumptions.

SM6. The Value of α in Eq. (2.12). As we argued in the main text, any pair (\mathbf{H}, \mathbf{S}) obeys the geometric mean bound (2.12) with $\alpha = 1/2$ and $\gamma = \max |\Lambda(\mathbf{H}, \mathbf{S})|$. In this section, we present numerical evidence that (2.12) often holds with $\alpha = 1/4$ and $\gamma \approx \max |\Lambda(\mathbf{H}, \mathbf{S})|$ for QSD problem instances, a substantial improvement on the provable bound. This $\alpha = 1/4$ behavior remains somewhat mysterious to us, and we have yet to discover a convincing explanation for why this behavior emerges.

The numerical validity of (2.12) with $\alpha = 1/4$ and $\gamma \approx \max |\Lambda(\mathbf{H}, \mathbf{S})|$ is demonstrated in Figure SM5. In these plots, we plot $y = |\mathbf{v}_j^* \mathbf{H} \mathbf{v}_k| / \max(\lambda_j, \lambda_k)$ against $x = \min(\lambda_j, \lambda_k) / \max(\lambda_j, \lambda_k)$ over all indices j and k for several different QSD instances, where we use the notation from section 2.2 that $(\lambda_j, \mathbf{v}_j)$ represents the j th largest eigenpair of \mathbf{S} . Since the accurately computable eigenvalues span a range roughly on the order of the inverse machine precision ($\approx 10^{16}$ in double precision), we only plot pairs (x, y) corresponding to indices j and k for which $\min(\lambda_j, \lambda_k) \geq 10^{-16} \lambda_1$. The bound (2.12) holds only if all points (x, y) (as well as those numerically incomputable) lie below a power law curve $y \leq \gamma x^{1-\alpha}$. The curve $\max |\Lambda(\mathbf{H}, \mathbf{S})| \cdot x^{0.75}$ is shown on each of the subplots in Figure SM5, and it lies above almost all of the pairs (x, y) . We consider this convincing evidence of the validity of (2.12) with $\alpha = 1/4$ and $\gamma \approx \max |\Lambda(\mathbf{H}, \mathbf{S})|$ for the QSD examples we tested. We suspect this relation will continue to hold for “reasonable” QSD instances, though we lack a precise definition of “reasonable” and a formal argument justifying this suspicion.

SM7. Extra Figures. Finally, we conclude with some additional figures concerning additional numerical experiments. Figure SM6 shows the smallest eigenvalue

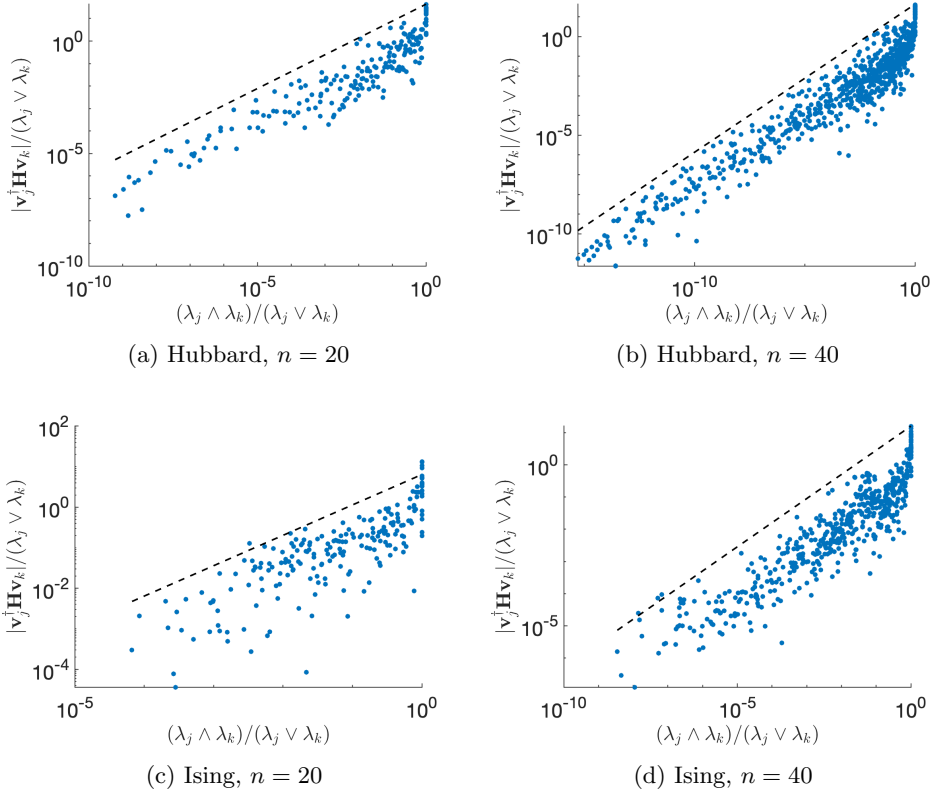


FIG. SM5. Scatter plot of values $y = |v_j^* H v_k|/(\lambda_j \vee \lambda_k)$ versus $x = \min(\lambda_j, \lambda_k)/\max(\lambda_j, \lambda_k)$ over all indices $j, k = 1, \dots, n$ for which $\min(\lambda_j, \lambda_k) \geq 10^{-16} \lambda_1$ for Hubbard model (top) and Ising model (bottom) with $n = 20$ (left) and 40 (right). Shown as a dashed black line is $\max |\Lambda(H, S)| \cdot x^{3/4}$, demonstrating that (2.12) holds numerically with $\alpha = 1/4$ and $\gamma \approx \max |\Lambda(H, S)|$.

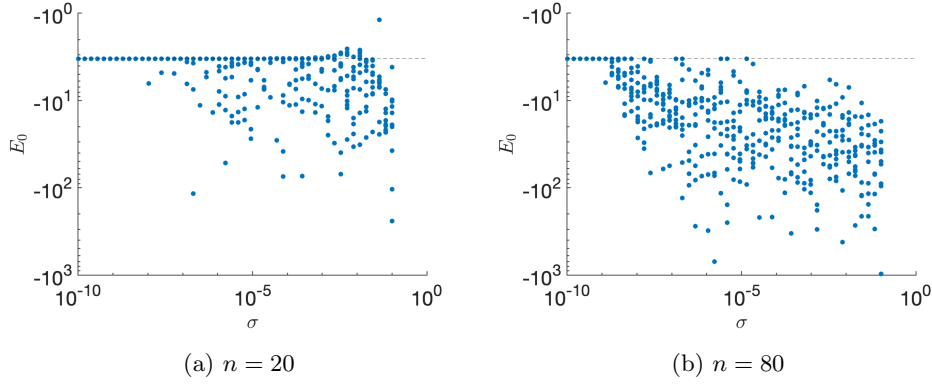


FIG. SM6. Least eigenvalues computed from the perturbed pair (\tilde{H}, \tilde{S}) with a fixed threshold $10^{-8} \|\mathbf{S}\|$. Shown are 10 random initializations of the noise for several random noise levels σ for the Hubbard example with $n = 20$ (left) and $n = 80$ (right). The true eigenvalue is shown for reference as a horizontal dashed line.

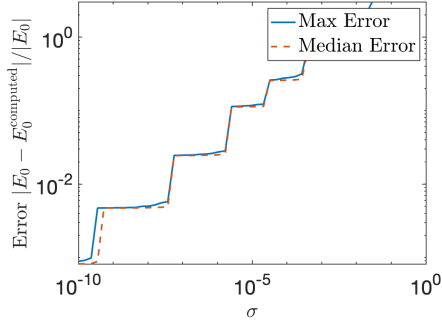
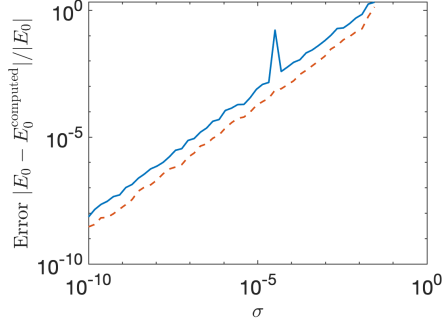
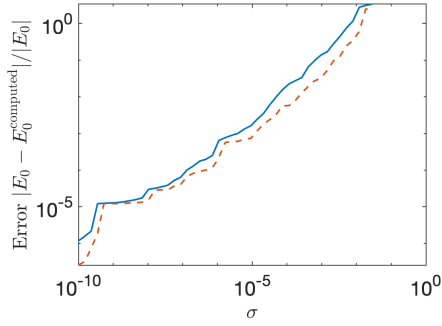
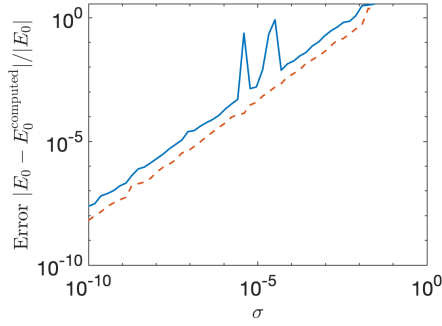
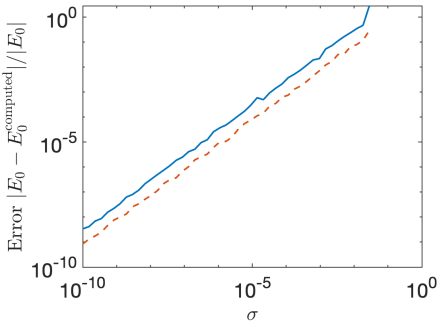
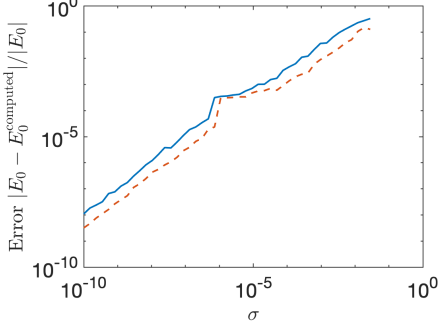
(a) Hubbard, $L = 10$, $U = 8$, $n = 10$ (b) Hubbard, $L = 10$, $U = 8$, $n = 80$ (c) Hubbard, $L = 10$, $U = 10$, $n = 30$ (d) Hubbard, $L = 10$, $U = 10$, $n = 80$ (e) Hubbard, $L = 6$, $U = 8$, $n = 40$ (f) Ising, $L = 8$, $n = 20$

FIG. SM7. Maximum (blue solid) and median (red dashed) error over 100 initializations for eigenvalues computed from the noise-perturbed pair $(\tilde{\mathbf{H}}, \tilde{\mathbf{S}})$ using thresholding with threshold parameter $25\sigma\|\tilde{\mathbf{S}}\|$ for Hubbard and Ising models for various parameters not considered in Figure 3.

computed when a fixed threshold parameter is used, independent of the noise level. Figures SM7, SM8, and SM9 provide more parameter settings for the Hubbard and Ising models for the Figures 3, 4, and SM2.

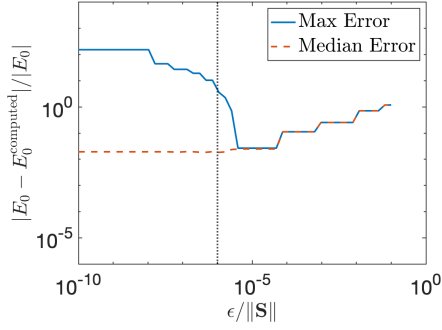
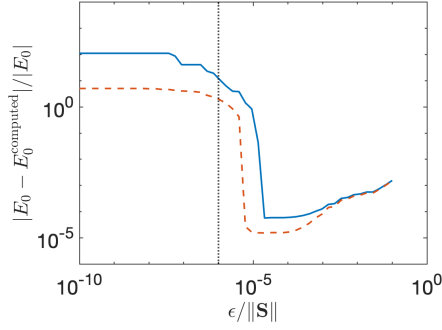
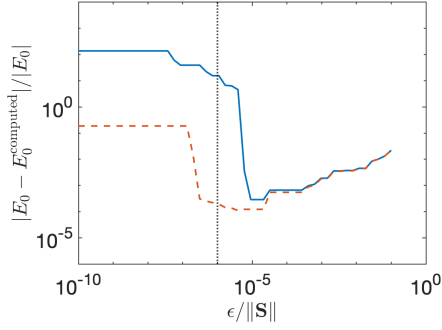
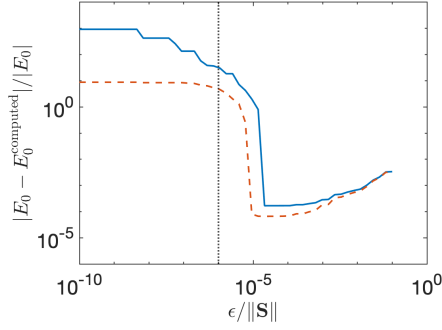
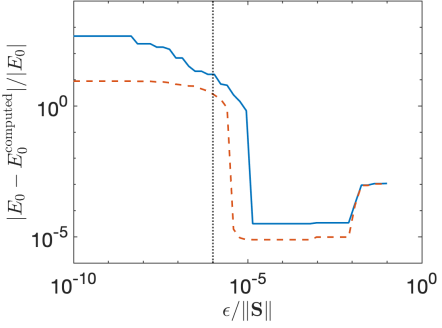
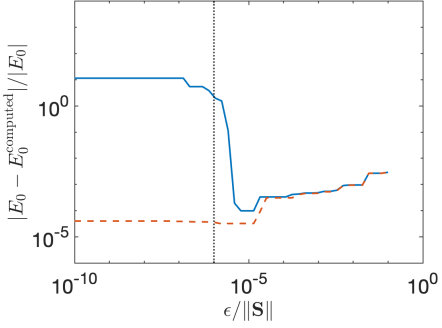
(a) Hubbard, $L = 10$, $U = 8$, $n = 10$ (b) Hubbard, $L = 10$, $U = 8$, $n = 80$ (c) Hubbard, $L = 10$, $U = 10$, $n = 30$ (d) Hubbard, $L = 10$, $U = 10$, $n = 80$ (e) Hubbard, $L = 6$, $U = 8$, $n = 40$ (f) Ising, $L = 8$, $n = 20$

FIG. SM8. Maximum (blue solid) and median (red dashed) error over 100 initializations for eigenvalues computed from the noise-perturbed pair $(\mathbf{H}, \tilde{\mathbf{S}})$ using thresholding for various values of the threshold ϵ for a fixed noise level $\sigma = 10^{-6}$ (dotted black line) for Hubbard and Ising models for various parameters not considered in Figure 4.

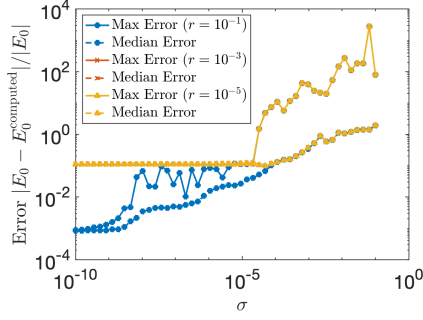
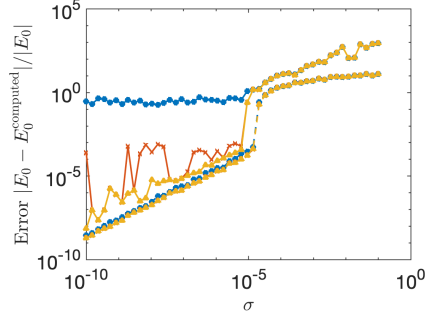
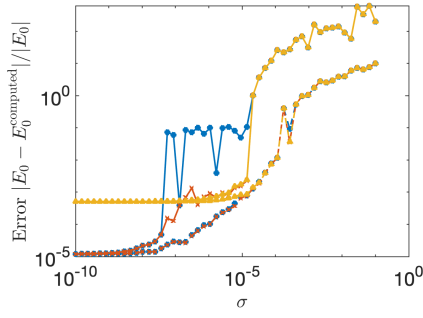
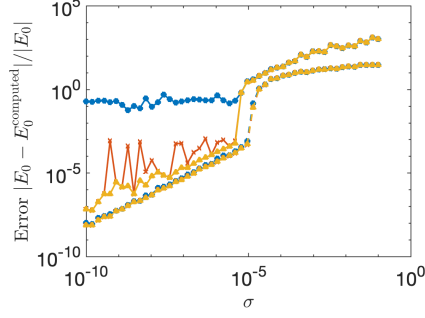
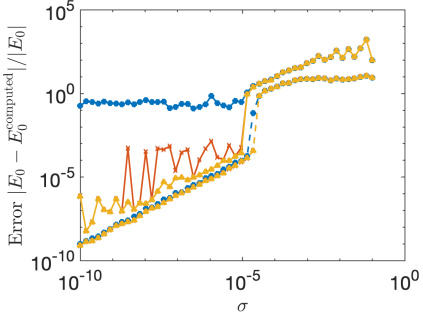
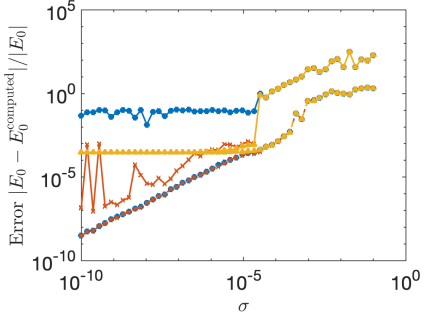
(a) Hubbard, $L = 10$, $U = 8$, $n = 10$ (b) Hubbard, $L = 10$, $U = 8$, $n = 80$ (c) Hubbard, $L = 10$, $U = 10$, $n = 30$ (d) Hubbard, $L = 10$, $U = 10$, $n = 80$ (e) Hubbard, $L = 6$, $U = 8$, $n = 40$ (f) Ising, $L = 8$, $n = 20$

FIG. SM9. Maximum and median error over 100 initializations for eigenvalues computed from the noise-perturbed pair $(\tilde{\mathbf{H}}, \tilde{\mathbf{S}})$ using automatically tuned thresholding (Algorithm SM3.1) for three cutoffs $r \in \{10^{-1}, 10^{-3}, 10^{-5}\}$ for various values of the noise level σ for Hubbard and Ising models for various parameters not considered in Figure SM2.

REFERENCES

- [SM1] R. BHATIA, *Matrix Analysis*, vol. 169 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1997, <https://doi.org/10.1007/978-1-4612-0653-8>.
- [SM2] B. N. PARLETT, *The Symmetric Eigenvalue Problem*, Classics in Applied Mathematics, SIAM, 1998, <https://doi.org/10.1137/1.9781611971163>.
- [SM3] G. W. STEWART AND J.-G. SUN, *Matrix Perturbation Theory*, Computer Science and Scientific Computing, Academic Press, 1st edition ed., 1990.