

Ground and Excited States from Ensemble Variational Principles

Lexin Ding^{1,2}, Cheng-Lin Hong^{1,2}, and Christian Schilling^{1,2}

¹Faculty of Physics, Arnold Sommerfeld Centre for Theoretical Physics (ASC),
Ludwig-Maximilians-Universität München, Theresienstr. 37, 80333 München, Germany

²Munich Center for Quantum Science and Technology (MCQST), Schellingstrasse 4, 80799 München, Germany

The extension of the Rayleigh-Ritz variational principle to ensemble states $\rho_w \equiv \sum_k w_k |\Psi_k\rangle\langle\Psi_k|$ with fixed weights w_k lies ultimately at the heart of several recent methodological developments for targeting excitation energies by variational means. Prominent examples are density and density matrix functional theory, Monte Carlo sampling, state-average complete active space self-consistent field methods and variational quantum eigensolvers. In order to provide a sound basis for all these methods and to improve their current implementations, we prove the validity of the underlying critical hypothesis: Whenever the ensemble energy is well-converged, the same holds true for the ensemble state ρ_w as well as the individual eigenstates $|\Psi_k\rangle$ and eigenenergies E_k . To be more specific, we derive linear bounds $d_- \Delta E_w \leq \Delta Q \leq d_+ \Delta E_w$ on the errors ΔQ of these sought-after quantities. A subsequent analytical analysis and numerical illustration proves the tightness of our universal inequalities. Our results and particularly the explicit form of $d_{\pm} \equiv d_{\pm}^{(Q)}(w, E)$ provide valuable insights into the optimal choice of the auxiliary weights w_k in practical applications.

1 Introduction

Quantum excitations play a crucial role in the understanding of quantum systems in general, and they lie at the heart of numerous applications in physics, chemistry, materials science and biology [1–5]. Unlike ground states, which can nowadays be numerically calculated with

high accuracy and ever-increasing efficiency [6, 7], the study of excited-state properties is still a long-standing computational challenge [8–15]. The complexity of the latter lies in the multi-configurational character of the corresponding wave functions [16], as well as in the lack of a variational principle for targeting individual excited states exclusively.

As far as the computation of ground states is concerned, the effectiveness of numerous numerical methods [17–19] rests on the underlying Rayleigh-Ritz (RR) variational principle: For any Hamiltonian

$$\hat{H} \equiv \sum_{k=0}^{D-1} E_k |\Psi_k\rangle\langle\Psi_k| \quad (1)$$

on a D -dimensional Hilbert space $\mathcal{H}$ the ground state energy E_0 can be obtained by variational means according to

$$E_0 \leq \langle\Psi|\hat{H}|\Psi\rangle, \quad \forall |\Psi\rangle \in \mathcal{H}. \quad (2)$$

The extremality within the energy spectrum of the sought-after minimal eigenvalue E_0 of $\hat{H}$ is absolutely crucial for the effectiveness and predictive power of variational ground state approaches. It namely implies (see, e.g., Ref. [20] and Sec. 2.1) that the accuracy of a trial state $|\Psi\rangle$ increases as the variational energy $\langle\Psi|\hat{H}|\Psi\rangle$ approaches the ground state energy E_0 , and equality with the exact ground state is reached if and only if the variational energy equals E_0 .

In order to target the low-lying excitation spectrum by variational means, Gross, Oliveira and Kohn [21] extended the RR variational principle to ensemble states $\tilde{\rho}_w \equiv \sum_{k=0}^{K-1} w_k |\tilde{\Psi}_k\rangle\langle\tilde{\Psi}_k|$ with

Christian Schilling: c.schilling@lmu.de

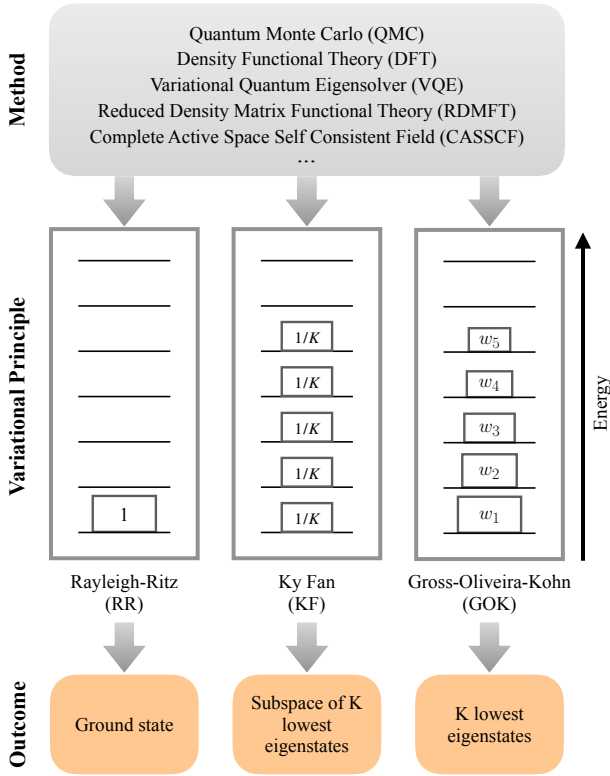


Figure 1: Illustration of the Rayleigh-Ritz (RR), Ky Fan (KF), and Gross-Oliveira-Kohn (GOK) variational principles and their scope in the context of recent methodological developments of ground and excited state approaches.

fixed spectrum $\mathbf{w}$:

$$\begin{aligned}
 E_{\mathbf{w}} &\equiv \sum_{k=0}^{K-1} w_k E_k \leq \text{Tr}[\tilde{\rho}_{\mathbf{w}} \hat{H}] \\
 &= \sum_{k=0}^{K-1} w_k \langle \tilde{\Psi}_k | \hat{H} | \tilde{\Psi}_k \rangle.
 \end{aligned} \tag{3}$$

The first K eigenenergies E_k and eigenstates $|\Psi_k\rangle$ of $\hat{H}$ are then obtained from the spectral decomposition of the resulting minimizer $\rho_{\mathbf{w}} \equiv \sum_{k=0}^{K-1} w_k |\Psi_k\rangle \langle \Psi_k|$.

This Gross-Oliveira-Kohn (GOK) variational principle naturally extends the scope of ground state methods to excited states (see the overview in Figure 1), and has recently become a key ingredient for the conceptual and numerical developments in quantum sciences due to a number of reasons: First, since the common approach in density functional theory (DFT) [22] to target excited states through a time-dependent formulation [23, 24] suffers from conceptual and practical limitations, ensemble/GOK-DFT [25–37] has emerged as a promising alternative. Second, due

to the difficulties of DFT in dealing with strong correlation effects, a one-particle reduced density matrix functional theory (RDMFT) [38–41] based on the GOK variational principle for targeting selected eigenstates has been proposed in 2021 ($\mathbf{w}$ -RDMFT) [42–46]. Third, wave function-based approaches that calculate the lowest K eigenstates consecutively suffer from an uncontrollable accumulation of errors. Utilizing instead the GOK variational principle with flexible weights $\mathbf{w}$ provides a more direct approach to the specific eigenstates of interest, which is successfully exploited, e.g., in quantum Monte Carlo techniques [47–49], some traditional quantum chemical methods [50–54] and quantum computing [55–63].

However, despite its broad applicability in recent years, two crucial aspects regarding the GOK variational principle have not been properly formulated nor addressed yet. (i) There are two seemingly misaligned objectives which first need to be reconciled: It is the ensemble energy that is minimized, yet the interest lies in the individual eigenstates and their energies. Although reaching the exact ensemble energy $E_{\mathbf{w}}$ implies that the eigenstates of the ensemble are exact as well [21], it is from a mathematical point of view not clear at all that the GOK variational principle maintains this predictive power for the practically relevant case of *imperfect* convergence. Indeed, particularly for quasidegenerate weights w_k , one could easily imagine an incorrect ensemble $\tilde{\rho}_{\mathbf{w}} = \sum_{k \geq 0} w_k |\tilde{\Psi}_k\rangle \langle \tilde{\Psi}_k| \neq \rho_{\mathbf{w}}$ which, however, (almost) attains the exact energy $E_{\mathbf{w}}$ due to a cancellation of individual energy errors $\tilde{E}_k - E_k$. (ii) There is an under-explored degree of freedom in the choice of $\mathbf{w}$. The validity of the GOK variational principle is independent of the specific choices of weights. Yet clearly, a variational algorithm can be biased towards a specific eigenstate, if it is assigned a larger weight. One extreme example would be the reduction of the GOK variational principle to the original RR version, or the one of equal weights, when virtually no information regarding the individual eigenstates could be gained. It is the ultimate goal of our work to address both of the above issues comprehensively and settle them conclusively.

The paper is structured as follows. First, we discuss the GOK variational principle and point out related caveats in Sec. 2.1, formulate con-

cisely our scientific problem in Sec. 2.2 and introduce in Sec. 2.3 two elegant concepts from convex geometry required for its solution. In Sec. 3, we present our main results, namely universal bounds, $d_- \Delta E_{\mathbf{w}} \leq \Delta Q \leq d_+ \Delta E_{\mathbf{w}}$, on the error ΔQ of the variational ensemble state, the eigenstates and eigenenergies. In Sec. 4, we then confirm by analytically and numerically means the tightness of these bounds and identify optimal choices of the weights for practical applications based on the explicit form of $d_{\pm} \equiv d_{\pm}^{(Q)}(\mathbf{w}, \mathbf{E})$. Last but not least, in Sec. 5, as a proof of concept, we illustrate our results in the context of variational quantum eigensolvers.

2 Notations, Concepts and Scientific Problem

2.1 GOK variational principle

In this section we provide a brief recap of the GOK variational principle, identify some crucial caveats in its practical applications and explain what is required from a variational principle in order to have a sound basis.

Attempts to extend the RR ground state variational principle date back to the late 40s, when Kohn and Ky Fan (the latter provided a full proof) established that the energy of an equiensemble state of rank K is bounded from above by the sum of the K lowest energy eigenvalues [64, 65]:

Theorem 1 (Ky Fan (1949)). *For any Hamiltonian (1) and any integer $K \leq D$ the inequality*

$$\sum_{k=0}^{K-1} E_k^{\uparrow} \leq \sum_{k=0}^{K-1} \langle \tilde{\Psi}_k | \hat{H} | \tilde{\Psi}_k \rangle \quad (4)$$

holds for any set of K orthonormal states $\{|\tilde{\Psi}_k\rangle\}_{k=0}^{K-1}$. Moreover, as long as $E_K > E_{K-1}$, equality is attained if and only if the span of the latter coincides with the span of the first K eigenstates $\{|\Psi_k\rangle\}_{k=0}^{K-1}$.

Here and in the following, an arrow $\uparrow$ ($\downarrow$) indicates that the entries of a vector are arranged increasingly (decreasingly).

It was almost 40 years later when Gross, Oliveira and Kohn [21] further extended Theorem 1 to ensembles with arbitrary weights:

Theorem 2 (Gross, Oliveira, Kohn (1988)). *For any Hamiltonian (1) and any probability distribution $\mathbf{w} \in \mathbb{R}^D$ ($w_i \geq 0$ and $\sum_{i=0}^{D-1} w_i = 1$), the inequality*

$$E_{\mathbf{w}} \equiv \sum_{k=0}^{D-1} w_k^{\downarrow} E_k^{\uparrow} \leq \text{Tr}[\tilde{\rho}_{\mathbf{w}} \hat{H}] = \sum_{k=0}^{D-1} w_k \langle \tilde{\Psi}_k | \hat{H} | \tilde{\Psi}_k \rangle. \quad (5)$$

holds for any ensemble state $\tilde{\rho}_{\mathbf{w}} \equiv \sum_{k=0}^{D-1} w_k |\tilde{\Psi}_k\rangle\langle\tilde{\Psi}_k|$ with spectrum $\mathbf{w}$. Moreover, as long as the energy spectrum $\mathbf{E}$ of $\hat{H}$ is non-degenerate, equality in (5) is attained if and only if $\tilde{\rho}_{\mathbf{w}}$ and $\rho_{\mathbf{w}} \equiv \sum_{k=0}^{D-1} w_k^{\downarrow} |\Psi_k\rangle\langle\Psi_k|$, and thus their spectral decompositions, coincide.

A couple of comments are in order here concerning this GOK variational principle. First, the weighted energy average is nothing else than the energy expectation value of the exact $\mathbf{w}$ -ensemble state $\rho_{\mathbf{w}}$, i.e., $E_{\mathbf{w}} \equiv \sum_{k=0}^{D-1} w_k^{\downarrow} E_k^{\uparrow} = \text{Tr}[\rho_{\mathbf{w}} \hat{H}] = \sum_{k=0}^{D-1} w_k \langle \Psi_k | \hat{H} | \Psi_k \rangle$ and thus (5) could be recast as $\text{Tr}[(\tilde{\rho}_{\mathbf{w}} - \rho_{\mathbf{w}}) \hat{H}] \geq 0$ for all $\mathbf{w}$ -ensemble states $\tilde{\rho}_{\mathbf{w}}$. Second, the predictive power of the GOK variational principle depends on how degenerate the energy spectrum $\mathbf{E}$ is. To explain this, let us consider the (practically less interesting) case of a fully degenerate spectrum, i.e., $\hat{H} = E_0 \hat{\mathbb{1}}$. Any $\mathbf{w}$ -ensemble state would then saturate inequality (5). Accordingly, the information that a given $\tilde{\rho}_{\mathbf{w}}$ saturates (5) would not tell us anything about this ensemble state. On the contrary, for a non-degenerate spectrum $E_0 < E_1 < \dots < E_{D-1}$, the saturation of (5) by some $\tilde{\rho}_{\mathbf{w}}$ would imply that its eigenbasis $\{|\tilde{\Psi}_k\rangle\}_{k=0}^{D-1}$ coincides (up to irrelevant phase factors) with the eigenbasis $\{|\Psi_k\rangle\}_{k=0}^{D-1}$ of $\hat{H}$ and $\rho_{\mathbf{w}}$, respectively. Given the goal of our paper, namely to elaborate on and confirm the predictive power of the GOK variational principle, this comment would suggest at first sight to study all scenarios of different energy multiplicities. Yet, there are good reasons to restrict our work to non-degenerate energy spectra only: In practice, degeneracies typically emerge from symmetries of the Hamiltonian, which translates into a block-diagonal structure, $\hat{H} \equiv \oplus_Q \hat{H}_Q$, with respect to the different symmetry sectors. One could then apply the entire reasoning of our work to each non-degenerate block/symmetry sector separately. Moreover, as our main results in Sec. 3 will confirm, the case of degeneracies is often either already covered as a suitable limit of the non-degenerate case, or alter-

natively it just does not make sense anymore to elaborate on the predictive power of GOK (as the example $\hat{H} = E_0 \hat{1}$ above illustrated). Third, in applications of the GOK variational principle to realistic quantum-many body systems one is typically interested only in the lowest few (K) eigenstates in the exponentially large Hilbert space. Hence, one then considers only $\mathbf{w}$ -vectors with K non-vanishing weights, $w_0 > w_1 > \dots > w_{K-1} > w_K = \dots = w_{D-1} = 0$. Of course, this in turn implies that the predictive power of the GOK variational principle then restricts to the lowest K eigenenergies and eigenstates, and that possible degeneracies of the higher energy levels E_k , $k \geq K$, will not affect our derivation and results anymore. Fourth, while a brute-force proof of Theorem 2 was given in Ref. [21], we will provide a particularly compact and simple one in Sec. 2.3, after having introduced some effective tools from convex analysis.

According to Theorem 2, the ensemble energy $\tilde{E}_{\mathbf{w}}$ of any $\tilde{\rho}_{\mathbf{w}} \equiv \sum_{k=0}^{D-1} w_k |\tilde{\Psi}_k\rangle\langle\tilde{\Psi}_k|$ is an upper bound to the true minimum $E_{\mathbf{w}}$. This, however, does not establish yet a sound basis for a variational principle that can be used for calculating individual eigenstates and their energies. Indeed, when the minimum $E_{\mathbf{w}}$ is not exactly attained, Theorem 2 does not make any predictions about the meaning or quality of the individual states $|\tilde{\Psi}_k\rangle$. That is quite in contrast to ground state calculations through the RR variational principle: It is well understood that for any pure trial state $|\Psi\rangle\langle\Psi|$, its Hilbert-Schmidt distance to the exact ground state $|\Psi_0\rangle\langle\Psi_0|$ ($\|\hat{A}\|_{\text{HS}} \equiv \sqrt{\text{Tr}[\hat{A}^\dagger \hat{A}]}$),

$$\Delta\Psi_0 \equiv \frac{1}{2} \left\| |\Psi\rangle\langle\Psi| - |\Psi_0\rangle\langle\Psi_0| \right\|_{\text{HS}}^2 = 1 - |\langle\Psi|\Psi_0\rangle|^2, \quad (6)$$

is linearly bounded from below and above by the variational error $\Delta E_0 \equiv \langle\Psi|\hat{H}|\Psi\rangle - E_0 \geq 0$ according to (see, e.g., Ref. [20])

$$q_- \Delta E_0 \leq \Delta\Psi_0 \leq q_+ \Delta E_0, \quad (7)$$

where $q_- \equiv 1/(E_{D-1} - E_0)$ and $q_+ \equiv 1/(E_1 - E_0)$. This means nothing else than that the variational state is exact if and only if the energy is fully converged and otherwise the errors $\Delta E_0, \Delta\Psi_0$ are related through linear bounds. It is exactly this *predictive power* that makes the RR variational principle an effective and meaningful variational principle. It is therefore the ultimate goal of our work to elaborate on and eventually confirm such

predictive power also for the GOK variational principle. This means to understand in quantitative terms whether, when and how the ensemble error $\Delta E_{\mathbf{w}}$ bounds the errors of (i) the ensemble state $\rho_{\mathbf{w}}$, (ii) the individual eigenstates $|\Psi_k\rangle$, and (iii) the individual eigenenergies E_k . This central task of our work can become quite delicate — in contrast to the context of the RR variational principle — as the following basic example illustrates:

An instructive example

Consider a three-level system with Hamiltonian $\hat{H} = -|0\rangle\langle 0| + |2\rangle\langle 2|$ and choose the ensemble weights $\mathbf{w} = (1/2, 1/2, 0)$. In this case both density operators $\rho_{\mathbf{w}} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$ and $\tilde{\rho}_{\mathbf{w}} = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$, where $|\pm\rangle \equiv 1/\sqrt{2}(|0\rangle \pm |1\rangle)$, are minimizers of the ensemble energy (both states are actually the same). However, the states $|\pm\rangle$ are quite different from the first two eigenstates $|0\rangle, |1\rangle$ of $\hat{H}$ and accordingly have large errors in their energies.

While the negative consequences of a degenerate weight vector $\mathbf{w}$ were actually rather obvious, this basic example clearly reveals the depth of the scientific problem at hand: the anticipated error bounds must depend in a non-trivial way on (the gaps between) the weights w_k , and in particular whenever two weights become identical some bounds must become meaningless.

2.2 The scientific problem

In this section, we formulate our scientific problem in concise terms. For this, we first assume a fixed non-degenerate Hamiltonian $\hat{H}$ on a D -dimensional complex Hilbert space $\mathcal{H}$ according to (1) and denote its vector of increasingly ordered energy eigenvalues by $\mathbf{E} \equiv \mathbf{E}^\uparrow$. For each choice of a spectral weight vector $\mathbf{w} \equiv \mathbf{w}^\downarrow$, we then define the corresponding manifold of density operators with spectrum $\mathbf{w}$ to which GOK's variational principle, Theorem 2, refers to:

$$\mathcal{D}_{\mathbf{w}} \equiv \left\{ \tilde{\rho}_{\mathbf{w}} = \sum_{k=0}^{D-1} w_k |\tilde{\Psi}_k\rangle\langle\tilde{\Psi}_k| \left| \langle\tilde{\Psi}_k|\tilde{\Psi}_l\rangle = \delta_{kl} \right. \right\}. \quad (8)$$

It is worth noticing that the manifold $\mathcal{D}_{\mathbf{w}}$ is identical to the set of all unitary conjugations of the

exact ensemble state $\rho_{\mathbf{w}} \equiv \sum_{k=0}^{D-1} w_k |\Psi_k\rangle\langle\Psi_k|$, i.e.,

$$\mathcal{D}_{\mathbf{w}} = \{\hat{U}\rho_{\mathbf{w}}\hat{U}^\dagger \mid \hat{U} \text{ a unitary on } \mathcal{H}\} \quad (9)$$

and accordingly every element $\tilde{\rho}_{\mathbf{w}}$ can be expressed as

$$\tilde{\rho}_{\mathbf{w}}(U) = \hat{U}\rho_{\mathbf{w}}\hat{U}^\dagger \quad (10)$$

for some unitary $\hat{U}$. To each quantum state $\tilde{\rho}_{\mathbf{w}}$ in $\mathcal{D}_{\mathbf{w}}$ we can associate the corresponding vector $\tilde{\mathbf{E}} \equiv (\tilde{E}_k)_{k=0}^{D-1}$ of energy expectation values $\tilde{E}_k = \langle\tilde{\Psi}_k|\hat{H}|\tilde{\Psi}_k\rangle$ with respect to $\tilde{\rho}_{\mathbf{w}}$'s eigenbasis $\{|\tilde{\Psi}_k\rangle\}_{k=0}^{D-1}$.

Elaborating on the predictive power of the GOK variational principle means nothing else than to predict the errors of physical properties of variational states $\tilde{\rho}_{\mathbf{w}} \in \mathcal{D}_{\mathbf{w}}$ as function of their ensemble energy error

$$\Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}}) \equiv \text{Tr}[(\tilde{\rho}_{\mathbf{w}} - \rho_{\mathbf{w}})\hat{H}]. \quad (11)$$

Due to its relevance for numerical calculations in practice, a particular emphasis will lie on the regime where the ensemble energy error becomes small. With these definitions at hand, we can now formulate in concise terms the scientific problem that we are going to solve:

Scientific Problem: Predictive Power of GOK

Consider the GOK variational principle (Theorem 2) for a setting $(\mathcal{H}, \hat{H}, \mathbf{w})$. Determine for any relevant quantity $Q \equiv Q(\tilde{\rho}_{\mathbf{w}})$ the range

$$d_-^{(Q)}(\delta, \mathbf{w}, \hat{H}) \leq \Delta Q(\tilde{\rho}_{\mathbf{w}}) \leq d_+^{(Q)}(\delta, \mathbf{w}, \hat{H}), \quad (12)$$

of its possible errors $\Delta Q(\tilde{\rho}_{\mathbf{w}}) \equiv Q(\tilde{\rho}_{\mathbf{w}}) - Q(\rho_{\mathbf{w}})$ provided $\tilde{\rho}_{\mathbf{w}}$ has a (sufficiently small) energy error $\Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}}) = \delta > 0$. The relevant quantities are

- (i) the ensemble state $\rho_{\mathbf{w}}$
- (ii) all individual eigenstates $|\Psi_k\rangle$
- (iii) all individual eigenenergies E_k

A couple of comments are in order. First, the sought-after optimal bounds are given by

$$\begin{aligned} d_-^{(Q)}(\delta, \mathbf{w}, \hat{H}) &= \min_{\substack{\tilde{\rho}_{\mathbf{w}} \in \mathcal{D}_{\mathbf{w}} : \\ \Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}}) = \delta}} \Delta Q(\tilde{\rho}_{\mathbf{w}}) \\ d_+^{(Q)}(\delta, \mathbf{w}, \hat{H}) &= \max_{\substack{\tilde{\rho}_{\mathbf{w}} \in \mathcal{D}_{\mathbf{w}} : \\ \Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}}) = \delta}} \Delta Q(\tilde{\rho}_{\mathbf{w}}) \end{aligned} \quad (13)$$

Hence, it will be the challenge in the following to execute by analytical means the respective constrained minimization and maximization in (13) for all three relevant physical quantities. Moreover, it will become clear below what ‘ δ sufficiently small’ will mean. Finally, the crucial question will be how the bounds $d_{\pm}(\delta, \mathbf{w}, \hat{H})$ depend on δ . To anticipate our results, we are actually going to find that the bounds for all three relevant quantities depend *linearly* on the ensemble error $\Delta E_{\mathbf{w}} = \delta$. This in turn will confirm the predictive power of the GOK variational principle.

To execute analytically the optimizations in (13) we first need to understand how various relevant error quantities depend on $\tilde{\rho}_{\mathbf{w}}$ and $\hat{U}$ in (10), respectively. By using $\tilde{\rho}_{\mathbf{w}} \equiv \tilde{\rho}_{\mathbf{w}}(\hat{U})$, $\mathbf{X} \equiv (|U_{kl}|^2)_{k,l=0}^{D-1}$, $U_{kl} \equiv \langle\Psi_k|\hat{U}|\Psi_l\rangle = \langle\Psi_k|\tilde{\Psi}_l\rangle$, $\tilde{\mathbf{w}} \equiv \mathbf{X}\mathbf{w}$ and $\tilde{\mathbf{E}} \equiv \mathbf{X}^T\mathbf{E}$, rather straightforward calculations yield (one just needs to evaluate various traces in the eigenbasis of either $\tilde{\rho}_{\mathbf{w}}$ or $\rho_{\mathbf{w}}$ and $\hat{H}$, respectively)

$$\begin{aligned} \Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}}) &\equiv \text{Tr}[(\tilde{\rho}_{\mathbf{w}} - \rho_{\mathbf{w}})\hat{H}] \\ &= \mathbf{w} \cdot (\mathbf{X}^T - \mathbb{1})\mathbf{E} \\ &= \mathbf{w} \cdot (\tilde{\mathbf{E}} - \mathbf{E}) \\ &= (\tilde{\mathbf{w}} - \mathbf{w}) \cdot \mathbf{E} \end{aligned} \quad (14)$$

$$\begin{aligned} \Delta\rho_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}}) &\equiv \|\tilde{\rho}_{\mathbf{w}} - \rho_{\mathbf{w}}\|_{\text{HS}}^2 \\ &\equiv \text{Tr}[(\tilde{\rho}_{\mathbf{w}} - \rho_{\mathbf{w}})^2] \\ &= 2\mathbf{w} \cdot (\mathbf{w} - \mathbf{X}\mathbf{w}) \\ &\equiv 2\mathbf{w} \cdot (\mathbf{w} - \tilde{\mathbf{w}}), \end{aligned} \quad (15)$$

$$\begin{aligned} \Delta\Psi_k(\tilde{\rho}_{\mathbf{w}}) &\equiv \frac{1}{2} \|\tilde{\Psi}_k\langle\tilde{\Psi}_k| - |\Psi_k\rangle\langle\Psi_k|\|_{\text{HS}}^2 \\ &= 1 - |\langle\Psi_k|\tilde{\Psi}_k\rangle|^2 \\ &= 1 - X_{kk}, \end{aligned} \quad (16)$$

$$\begin{aligned} \Delta E_k(\tilde{\rho}_{\mathbf{w}}) &\equiv \text{Tr}[(\tilde{\Psi}_k\langle\tilde{\Psi}_k| - |\Psi_k\rangle\langle\Psi_k|)\hat{H}] \\ &= (\mathbf{X}^T\mathbf{E})_k - E_k \\ &\equiv \tilde{E}_k - E_k, \end{aligned} \quad (17)$$

The far-reaching observation here is that all relevant quantities (14)-(17) can be expressed as *linear* functions of either the matrix $\mathbf{X} \equiv (|U_{kl}|^2)_{k,l=0}^{D-1}$, or the even simpler vectors $\tilde{\mathbf{w}} = \mathbf{X}\mathbf{w}$ and $\tilde{\mathbf{E}} = \mathbf{X}^T\mathbf{E}$, respectively. Accordingly, in order to exploit this simplifying structure we need to understand better the sets of such matrices $\mathbf{X}$ and such vectors $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{E}}$. This is accomplished in the following section.

2.3 Birkhoff polytopes and permutohedra

In the spirit of the last comment in the previous section, the problem (13) of ensemble minimization and maximization can be conveniently cast in the language of Birkhoff polytopes (and their projections, permutohedra). Birkhoff polytopes are central mathematical objects that have been widely employed in various fields of physics (see [66] and references therein). The rich mathematical structures of these polytopes will turn out to be instrumental for the derivation of our main results and the comprehensive solution of the scientific problem (see p. 4).

The so-called Birkhoff polytope $\mathcal{B}_D$ of order D is a convex polytope consisting of D -by- D doubly stochastic matrices [67]. A matrix $\mathbf{X}$ is called doubly stochastic (or bistochastic) if its columns and rows consist of real, non-negative entries which for each row and column sum to 1,

$$\sum_{l=0}^{D-1} X_{lk} = \sum_{l=0}^{D-1} X_{kl} = 1, \quad \forall k = 0, 1, \dots, D-1. \quad (18)$$

Additionally, if there exists a unitary matrix $\mathbf{U}$ such that $X_{kl} = |U_{kl}|^2$, then $\mathbf{X}$ is called unistochastic. We denote the set of unistochastic matrices as $\mathcal{U}_D$ which according to Eqs. (14)-(17) will play a key role in the solution of the scientific problem. For $D \geq 3$, it turns out that the set $\mathcal{U}_D$ is no longer convex [66–68]. In particular, there exist doubly stochastic matrices that are not unistochastic, and therefore $\mathcal{U}_D$ is then a strict subset of $\mathcal{B}_D$.

In order to develop a better intuition for the Birkhoff polytope, we recall in particular the Birkhoff-von Neumann theorem [69, 70]: The Birkhoff polytope $\mathcal{B}_D$ is a convex polytope in $\mathbb{R}^{D \times D}$ of dimension $(D-1)^2$ whose extremal points are the permutation matrices [67, 71]. The latter are precisely those matrices which contain in each row and column one ‘1’ and $D-1$ ‘0’-entries, and together they form a discrete set with $D!$ elements denoted by $\mathcal{S}_D$. Accordingly, by definition, $\mathcal{B}_D$ follows as the convex hull $\text{Conv}(\mathcal{S}_D)$. Moreover, two extremal points (vertices) of $\mathcal{B}_D$ are connected by an edge if and only if they differ by a single cycle [72]. Since all permutation matrices are unistochastic, the convex hull $\text{Conv}(\mathcal{U}_D)$ is precisely $\mathcal{B}_D$. The relations among $\mathcal{B}_D$, $\mathcal{U}_D$ and $\mathcal{S}_D$ are depicted in the left panel of

Figure 2. They can be summarized as

$$\mathcal{S}_D \subseteq \mathcal{U}_D \subseteq \mathcal{B}_D, \quad (19)$$

and

$$\mathcal{B}_D = \text{Conv}(\mathcal{U}_D) = \text{Conv}(\mathcal{S}_D). \quad (20)$$

Since we need to minimize and maximize the linear functions (15)-(17) over a compact convex set according to Eq. (13), we discuss in the following some basics of linear optimization. We focus on minimization, yet analogous conclusions can be drawn also for maximization. First, we recall the general fact that real-valued *linear* functions $L(\mathbf{X})$ on compact sets attain their minima at extremal points. In particular, in our context this implies

$$\min_{\mathbf{X} \in \mathcal{U}_D} L(\mathbf{X}) = \min_{\mathbf{X} \in \mathcal{B}_D} L(\mathbf{X}) = \min_{\mathbf{P} \in \mathcal{S}_D} L(\mathbf{P}). \quad (21)$$

This becomes obvious in the geometric picture as it is illustrated in the left panel of Figure 2: to minimize $L(\mathbf{X})$ one needs to shift the blue hyperplane defined by $L(\mathbf{X}) = \text{const}$ in the direction of its normal vector until the boundary of the underlying set is reached. If we restrict, however, the minimization of $L(\mathbf{X})$ in (21) by one (or several) linear constraint $\mathcal{A}(\mathbf{X}) = c \in \mathbb{R}$ (in our case $\Delta E_{\mathbf{w}}(\mathbf{X}) = \delta$), in general only the following inequality holds

$$\min_{\mathbf{X} \in \mathcal{U}_D, \mathcal{A}(\mathbf{X})=c} L(\mathbf{X}) \geq \min_{\mathbf{X} \in \mathcal{B}_D, \mathcal{A}(\mathbf{X})=c} L(\mathbf{X}). \quad (22)$$

This crucial distinction between unconstrained and constrained minimization is illustrated in Figure 2. We can see that in the latter case (right panel), the minimizers of $L(\mathbf{X})$ over $\mathcal{B}_D$ and $\mathcal{U}_D$ do not necessarily coincide anymore.

In the particular case where $L(\mathbf{X}) \equiv l(\mathbf{X}\mathbf{w})$ with some potential constraint $\mathcal{A}(\mathbf{X}) \equiv a(\mathbf{X}\mathbf{w}) = c$ for some vector $\mathbf{w} \in \mathbb{R}^D$, the minimization of $L(\mathbf{X})$ over $\mathcal{U}_D$ and $\mathcal{B}_D$, respectively, can be simplified to a minimization of $l(\tilde{\mathbf{w}})$ over the simpler permutohedron $P(\mathbf{w})$ of $\mathbf{w}^1$, given by [73]

$$P(\mathbf{w}) \equiv \{\tilde{\mathbf{w}} \in \mathbb{R}^D \mid \exists \mathbf{X} \in \mathcal{B}_D : \tilde{\mathbf{w}} = \mathbf{X}\mathbf{w}\}. \quad (23)$$

The set $P(\mathbf{w})$ inherits the convex properties of $\mathcal{B}_D$, but its dimensionality is much lower. We

¹When $\mathbf{w}$ contains repeated elements, as in some of our cases where some entries of $\mathbf{w}$ can be zero, $P(\mathbf{w})$ is not in a strict sense a permutohedron. When $\mathbf{w}$ contains repeated elements, $P(\mathbf{w})$ no longer has the correct number of vertices and edges.

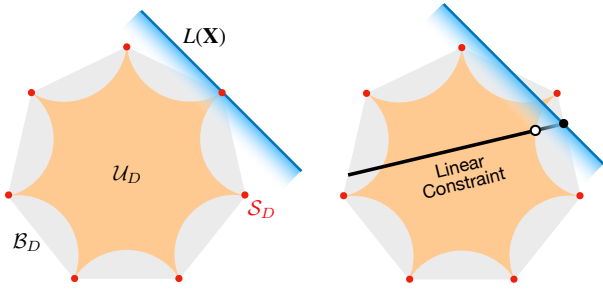


Figure 2: Schematic illustration of unconstrained (left) versus constrained minimization (right) of a linear function $L(\mathbf{X})$ (blue line) over the ‘grey’ Birkhoff polytope $\mathcal{B}_D$. The ‘orange’ subset $\mathcal{U}_D$ of unistochastic matrices and the discrete ‘red’ set $\mathcal{S}_D$ of extremal elements of $\mathcal{B}_D$ (permutation matrices) are also presented.

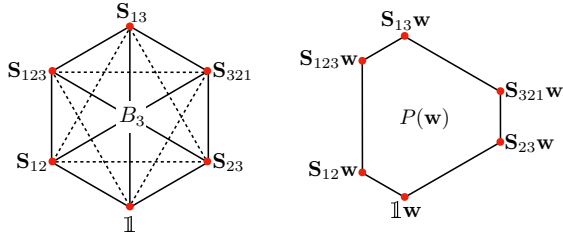


Figure 3: Schematic illustration of the four-dimensional Birkhoff polytope $\mathcal{B}_3$ (left) and the two-dimensional permutohedron $P(\mathbf{w})$ of a generic vector $\mathbf{w} \in \mathbb{R}^3$ (right). Vertices are presented as red dots and cyclic notation for the six permutations is used. Edges are depicted by solid or dashed lines. On the left panel, solid and dashed lines are used to indicate different lengths of edges.

present in Figure 3 a comparison between the Birkhoff polytope $\mathcal{B}_3$, and the permutohedron of a vector $\mathbf{w} \in \mathbb{R}^3$. Since the extremal points of $\mathcal{B}_3$ are permutation matrices, the extremal elements of $P(\mathbf{w})$ follow as the permutations of the vector $\mathbf{w}$. Since all permutations of the entries of a vector in three dimension are cyclic, all vertices of $\mathcal{B}_3$ are connected by an edge. However, in general, two extremal points (vertices) of $P(\mathbf{w})$ are connected if and only if they can be transformed into one another by an adjacent transposition $S_{i,i+1}$ [73]. Consequently, all edges of $P(\mathbf{w})$ consist of vectors of the form $\tilde{\mathbf{w}} = \mathbf{X}'\mathbf{w}$, where $\mathbf{X}'$ are unistochastic, and the following convex relaxation is exact

$$\begin{aligned} \min_{\mathbf{X} \in \mathcal{U}_D, \mathcal{A}(\mathbf{X})=c} L(\mathbf{X}) &= \min_{\mathbf{X} \in \mathcal{B}_D, \mathcal{A}(\mathbf{X})=c} L(\mathbf{X}) \\ &= \min_{\tilde{\mathbf{w}} \in P(\mathbf{w}), \mathcal{A}(\tilde{\mathbf{w}})=c} l(\tilde{\mathbf{w}}). \end{aligned} \quad (24)$$

We conclude this section by providing a proof of the GOK variational principle, Theorem 2. Its

simplicity, also compared to the original proof in Ref. [21], already demonstrates how effective the tools from convex analysis are that we have introduced in this section.

Proof. We know that $\text{Tr}[\tilde{\rho}_{\mathbf{w}}\hat{H}] = \mathbf{w}^T \mathbf{X}^T \mathbf{E}$ is linear in $\mathbf{X}$, where $\mathbf{X} \equiv (|\langle \Psi_k | \tilde{\Psi}_l \rangle|^2)_{k,l=0}^{D-1} \in \mathcal{U}_D$. Therefore, according to Eq. (21) and since $\mathcal{U}_D = \mathcal{U}_D^T$ is invariant under transposition, we can relax the problem of minimizing $\text{Tr}[\tilde{\rho}_{\mathbf{w}}\hat{H}]$ over the set $\mathcal{U}_D$ of unistochastic matrices $\mathbf{X}$ to the set $\mathcal{S}_D$ of permutation matrices $\mathbf{P}$. This yields

$$\begin{aligned} \min_{\tilde{\rho}_{\mathbf{w}}} \text{Tr}[\tilde{\rho}_{\mathbf{w}}\hat{H}] &= \min_{\mathbf{X} \in \mathcal{U}_D} \mathbf{w}^T \mathbf{X}^T \mathbf{E} \\ &= \min_{\mathbf{P} \in \mathcal{S}_D} \mathbf{w}^T \mathbf{P} \mathbf{E} \\ &= \mathbf{w}^\downarrow \cdot \mathbf{E}^\uparrow. \end{aligned} \quad (25)$$

In the last step we applied the rearrangement inequality [74]. $\square$

3 Error Bounds on quantum states and eigenenergies

In this section, we present our main results, namely according to the ‘Scientific Problem’ on p. 4 the lower and upper bounds (12) on the errors of the ensemble state, and the individual eigenstates and eigenenergies as function of the ensemble error (11).

First, we notice that the condition of a fixed error $\Delta E_{\mathbf{w}} = \delta > 0$ in the ensemble energy defines a linear constraint, which is geometrically represented as an hyperplane $\mathbb{A}(\delta)$ in $\mathcal{B}_D$, $P(\mathbf{w})$, and $P(\mathbf{E})$. Consequently, there is a straightforward strategy for determining the lower and upper bounds of the linear functions (15)-(17) according to (13): Following Sec. 2.3, one just needs to evaluate these functions at the vertices of the convex polytope obtained by intersecting $\mathbb{A}(\delta)$ with the corresponding original polytope $\mathcal{B}_D$, $P(\mathbf{w})$, and $P(\mathbf{E})$, respectively. The lowest/highest value that is found is nothing else than the sought-after bound $d_{\pm}^{(Q)}(\delta, \mathbf{w}, \hat{H})$ for the corresponding quantity Q .

In Figure 4 we illustrate our strategy for determining the bounds (12) according to (13). The hyperplane $\mathbb{A}(\delta)$ defined by $\Delta E_{\mathbf{w}} = \delta > 0$ intersects with the underlying polytope at the orange region. Since according to the GOK variational principle the ensemble error $\Delta E_{\mathbf{w}}$ is always non-negative we can distinguish the vertices of of $\mathcal{B}_D$

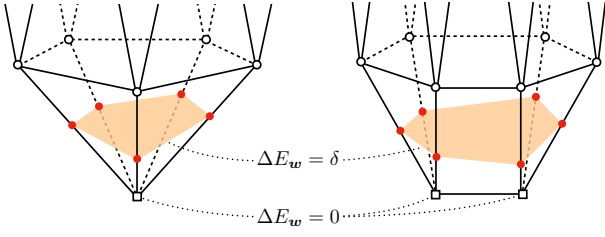


Figure 4: Linear constrained optimization on a polytope. Reference vertices ($\Delta E_{\mathbf{w}} = 0$) are represented as squares. The convex constraint domain ($\Delta E_{\mathbf{w}} = \delta$) is shown as the orange area, whose vertices are red circles. On the left panel there is only one reference vertex. On the right panel we show the case when there are multiple reference vertices.

and $P(\mathbf{w})$ and $P(\mathbf{E})$, respectively, into two categories: The so-called reference vertices characterized by a vanishing error, $\Delta E_{\mathbf{w}} = 0$, and the other ones whose error is finite, $\Delta E_{\mathbf{w}} > 0$. Since polytopes have finitely many vertices, we can always find a δ sufficiently small such that only reference vertices obey $\Delta E_{\mathbf{w}} < \delta$. When all entries in $\mathbf{w}$ are distinct, there will always be just one reference vertex: In the Birkhoff polytope $\mathcal{B}_D$ it is the identity matrix and in the permutohedra $P(\mathbf{w})$ and $P(\mathbf{E})$ the untransformed vectors $\mathbf{w}$ and $\mathbf{E}$, respectively. When only $K < D$ entries of $\mathbf{w}$ are positive and distinct (and the rest are zero), there will be several reference vertices. To understand this, we point out that permuting the energy levels above the K -th eigenstate does not change the ensemble energy.

By anticipating some insights from the following derivations, the criterion of δ sufficiently small can be made precise. For this, and for later purposes we introduce two auxiliary functions:

$$\begin{aligned} g_{\mathbf{w}, \mathbf{E}} &= \min_{\substack{k < D-1 \\ w_k > w_{k+1}}} (w_k^\downarrow - w_{k+1}^\downarrow)(E_{k+1}^\uparrow - E_k^\uparrow) \\ G_{\mathbf{w}, \mathbf{E}} &= (w_0^\downarrow - w_{D-1}^\downarrow)(E_{D-1}^\uparrow - E_0^\uparrow). \end{aligned} \quad (26)$$

The quantities $g_{\mathbf{w}, \mathbf{E}}$ and $G_{\mathbf{w}, \mathbf{E}}$ are precisely the minimal and maximal errors in the ensemble energy $E_{\mathbf{w}}$ when two eigenstates in the exact ensemble $\rho_{\mathbf{w}}$ are swapped. Moreover, our assumption of δ sufficiently small translates into the condition $\delta \leq g_{\mathbf{w}, \mathbf{E}}$. In that practically relevant regime, the derived bounds (13) will simplify to $d_{\pm}^{(Q)}(\delta, \mathbf{w}, \hat{H}) = d_{\pm}^{(Q)}(\mathbf{w}, \hat{H}) \delta$, as a consequence of the linearity of the constraint function (14) and target functions (15)-(17).

The procedure for determining $d_{\pm}^{(Q)}(\mathbf{w}, \hat{H})$ con-

sists of three steps: (i) Identification of all edges of the underlying polytope $\mathcal{B}_D$, $P(\mathbf{w})$ or $P(\mathbf{E})$ that intersects with $\mathbb{A}(\delta)$, (ii) computation of the vertices of the intersection (red vertices in Figure 4), and (iii) evaluation of the target function on all vertices of the intersection polytope and identification of its minimal and maximal value. In the following sections, we will determine $d_{\pm}^{(Q)}(\mathbf{w}, \hat{H})$ for two important classes of $\mathbf{w}$: (a) $\mathbf{w}$ with $w_0 > w_1 > \dots > w_{D-1} \geq 0$ which can be employed in practice only for smaller systems in order to compute the *entire* energy spectrum, and for larger systems (b) $\mathbf{w}$ with $w_0 > w_1 > \dots > w_K = w_{K+1} = \dots = w_{D-1} = 0$ for targeting the lowest K eigenstates.

3.1 Targeting the ensemble state

Acting in accordance with the ‘Scientific Problem’ on p. 4, the first quantity whose error range we quantify as function of the ensemble error $\Delta E_{\mathbf{w}}$ is the ensemble state $\rho_{\mathbf{w}}$. This leads to the GOK-analogue of the well-known estimate (7) for the Rayleigh-Ritz principle:

Theorem 3. *Consider the ‘Scientific Problem’ for a setting $(\mathcal{H}, \hat{H}, \mathbf{w})$ and recall (14), (15). The errors $\Delta \rho_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}})$ of the ensemble state and $\Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}})$ of the ensemble energy are universally related for all $\mathbf{w}$ -ensembles $\tilde{\rho}_{\mathbf{w}}$ according to*

$$a_{-}(\mathbf{w}, \mathbf{E}) \Delta E_{\mathbf{w}} \leq \Delta \rho_{\mathbf{w}} \leq a_{+}(\mathbf{w}, \mathbf{E}) \Delta E_{\mathbf{w}}, \quad (27)$$

where for $w_0 > w_1 > \dots > w_{D-1}$

$$\begin{aligned} a_{-}(\mathbf{w}, \mathbf{E}) &\equiv 2 \min_{k < D-1} \frac{w_k - w_{k+1}}{E_{k+1} - E_k} \\ a_{+}(\mathbf{w}, \mathbf{E}) &\equiv 2 \max_{k < D-1} \frac{w_k - w_{k+1}}{E_{k+1} - E_k}, \end{aligned} \quad (28)$$

and for $w_0 > w_1 > \dots > w_K = \dots = w_{D-1} = 0$ ($K < D-1$)

$$\begin{aligned} a_{-}(\mathbf{w}, \mathbf{E}) &\equiv 2 \min \left\{ \min_{k < K-1} \frac{w_k - w_{k+1}}{E_{k+1} - E_k}, \frac{w_{K-1}}{E_{D-1} - E_{K-1}} \right\} \\ a_{+}(\mathbf{w}, \mathbf{E}) &\equiv 2 \max_{k < K} \frac{w_k - w_{k+1}}{E_{k+1} - E_k}. \end{aligned}$$

Sketch of proof. As explained in Sec. 2.3, the specific linear form of (14) and (15) allows us to transform the underlying optimization problem (13) into one on the permutohedron $P(\mathbf{w})$. In

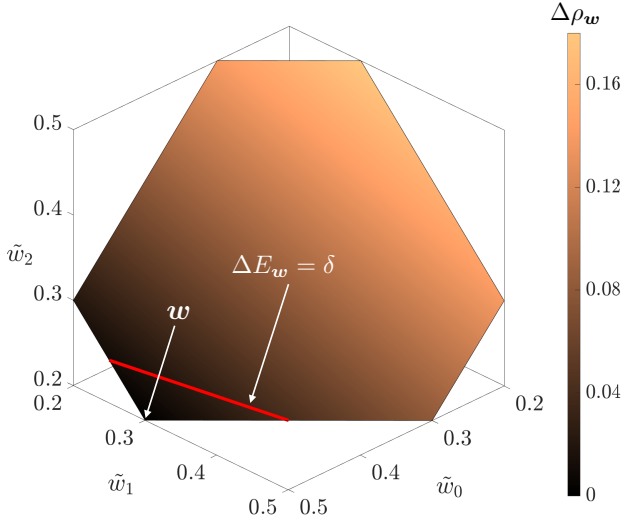


Figure 5: Permutohedron $P(\mathbf{w})$ of $\mathbf{w} = (0.5, 0.3, 0.2)$. The face color encodes the value of $\Delta\rho_{\mathbf{w}}(\tilde{\mathbf{w}})$ according to (15). The red line is the intersection between $P(\mathbf{w})$ and the hyperplane defined by $\Delta E_{\mathbf{w}}(\tilde{\mathbf{w}}) = \delta = 0.1$ where $\mathbf{E} = (-1, 0, 2)$.

the regime where $\Delta E_{\mathbf{w}} = \delta < g_{\mathbf{w}, \mathbf{E}}$, there is only one reference vertex, namely $\mathbf{w}$. If $\mathbf{w}$ is strictly decreasing ($K = D$), then we know from Sec. 2.3 that the only neighboring vertices are $\mathbf{S}_{i,i+1}\mathbf{w}$, and the connecting edges are convex combinations between $\mathbf{w}$ and $\mathbf{S}_{i,i+1}\mathbf{w}$. This is why the lower and upper bounds involve only differences between neighboring entries in $\mathbf{w}$ and $\mathbf{E}$. From here one can easily work out the intersection between $\Delta E_{\mathbf{w}} = \delta$ and $P(\mathbf{w})$. If $\mathbf{w}$ instead fulfills the second requirement ($K < D - 1$), then neighboring vertices to $\mathbf{w}$ include not only $\mathbf{S}_{i,i+1}\mathbf{w}$, but also $\mathbf{S}_{K-1,k+1}\mathbf{w}$ where $k \geq K$. The detailed proof of Theorem 3 is reserved to the Appendix.

We illustrate in Figure 5 the proof of Theorem 3 with an example Hamiltonian in three dimensions. For the concrete plot, we chose the energy spectrum $\mathbf{E} = (-1, 0, 2)$ and the ensemble weights $\mathbf{w} = (0.5, 0.3, 0.2)$. The hexagonal region in Figure 5 is the permutohedron $P(\mathbf{w})$ with elements $\tilde{\mathbf{w}}$. The red solid segment represents the intersection between $P(\mathbf{w})$ and the constraint $\Delta E_{\mathbf{w}} = \delta = 0.1$. The error of the ensemble state $\Delta\rho_{\mathbf{w}}$, which is encoded in color, has a global minimum at the reference vertex $\mathbf{w}$ (bottom left corner). The minimum and maximum of $\Delta\rho_{\mathbf{w}}$ within the red segment are obtained at the two end points, and the inequalities provided by Theorem 3 are explicitly saturated.

Knowing the accuracy of the ensemble state is

of central importance, as it allows us to make reliable prediction on the accuracy of the expectation value of any observable $\hat{A}$ with respect to the optimized state.

Corollary 1. *The error $\Delta A_{\mathbf{w}} \equiv \text{Tr}[(\tilde{\rho}_{\mathbf{w}} - \rho_{\mathbf{w}})\hat{A}]$ of the expectation of an observable $\hat{A}$ satisfies for all $\mathbf{w}$ -ensemble states $\tilde{\rho}_{\mathbf{w}}$*

$$|\Delta A_{\mathbf{w}}| \leq \|\hat{A}\|_{\text{HS}} \sqrt{a_+(\mathbf{w}, \mathbf{E}) \Delta E_{\mathbf{w}}}. \quad (29)$$

Proof. Since $\langle \mathbf{X}, \mathbf{Y} \rangle_{\text{HS}} \equiv \text{Tr}[\mathbf{X}^\dagger \mathbf{Y}]$ defines an inner product, the Cauchy-Schwarz inequality applies and yields $|\text{Tr}[\mathbf{X}^\dagger \mathbf{Y}]| \leq \|\mathbf{X}\|_{\text{HS}} \|\mathbf{Y}\|_{\text{HS}}$, where $\|\mathbf{X}\|_{\text{HS}} \equiv \sqrt{\langle \mathbf{X}, \mathbf{X} \rangle_{\text{HS}}}$. With $\mathbf{X} = \hat{A}$ and $\mathbf{Y} = \tilde{\rho}_{\mathbf{w}} - \rho_{\mathbf{w}}$, (29) then follows directly from Theorem 3. $\square$

3.2 Targeting the eigenstates

Next, we quantify the range of the errors of the individual eigenstates. We remark that for $k \geq K$ the error $\Delta\Psi_k$ cannot be estimated by $\Delta E_{\mathbf{w}}$ anymore since $|\tilde{\Psi}_k\rangle$ does not enter the ensemble $\tilde{\rho}_{\mathbf{w}}$ as a consequence of $w_{k \geq K} = 0$. For example, $\Delta\Psi_K$ can take any value between 0 and 1 even if $\Delta E_{\mathbf{w}} = 0$, as the energy of the K -th eigenstate does not contribute to the ensemble energy. Conversely, in order to extract accurate information about the target excited states, the corresponding weights must be non-zero and distinct from others.

Theorem 4. *Consider the ‘Scientific Problem’ for a setting $(\mathcal{H}, \hat{H}, \mathbf{w})$ and recall (14), (16). For any k with non-degenerate w_k , the error $\Delta\Psi_k(\tilde{\rho}_{\mathbf{w}})$ of the k -th eigenstate and $\Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}})$ of the ensemble energy are universally related for all $\mathbf{w}$ -ensembles $\tilde{\rho}_{\mathbf{w}}$ according to*

$$0 \leq \Delta\Psi_k \leq b_+^{(k)}(\mathbf{w}, \mathbf{E}) \Delta E_{\mathbf{w}}, \quad (30)$$

where

$$b_+^{(k)}(\mathbf{w}, \mathbf{E}) \equiv \begin{cases} \frac{1}{(w_0 - w_1)(E_1 - E_0)}, & k = 0 \\ \frac{1}{\min\{t_{k-1}, t_k\}}, & k > 0 \end{cases} \quad (31)$$

and $t_k \equiv (w_k - w_{k+1})(E_{k+1} - E_k)$.

Theorem 4 would directly yield a lower and upper bound on the sum $\Delta\Psi = \sum_{k=0}^{K-1} \Delta\Psi_k$ of various individual errors $\Delta\Psi_k$, yet they will not be tight. Instead, the optimal bounds on $\Delta\Psi$ is given by the following theorem. To formulate it

in a compact way and for subsequent considerations we recall the Heaviside step function $\Theta(x)$ defined as

$$\Theta(x) = \begin{cases} 1, & x > 0 \\ 1/2, & x = 0 \\ 0, & x < 0 \end{cases} \quad (32)$$

Theorem 5. Consider the ‘Scientific Problem’ for a setting $(\mathcal{H}, \hat{H}, \mathbf{w})$ and recall (14), (16). Then, the sum of the errors of the targeted eigenstates and the error $\Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}})$ of the ensemble energy are universally related for all $\mathbf{w}$ -ensembles $\tilde{\rho}_{\mathbf{w}}$ in the following way: When $w_0 > w_1 > \dots > w_{D-1}$, one has

$$\frac{2\Delta E_{\mathbf{w}}}{G_{\mathbf{w}, \mathbf{E}}} \leq \sum_{k=0}^{D-1} \Delta \Psi_k \leq \frac{2\Delta E_{\mathbf{w}}}{g_{\mathbf{w}, \mathbf{E}}}, \quad (33)$$

where $g_{\mathbf{w}, \mathbf{E}}$ and $G_{\mathbf{w}, \mathbf{E}}$ are defined in (26). When $w_0 > w_1 > \dots > w_K = \dots = w_{D-1} = 0$ for some $K < D-1$ one has

$$\begin{aligned} \frac{\Delta E_{\mathbf{w}}}{G_{\mathbf{w}, \mathbf{E}}} &\leq \sum_{k=0}^{K-1} \Delta \Psi_k \\ &\leq \max_{k < K} \left\{ \frac{2\Theta(K-k-1)}{(w_k - w_{k+1})(E_{k+1} - E_k)} \right\} \Delta E_{\mathbf{w}}. \end{aligned} \quad (34)$$

Since the error of an eigenstate given by (16) cannot be interpreted as a function of $\mathbf{X}\mathbf{w}$ or $\mathbf{X}^T \mathbf{E}$, the proofs of Theorems 4 and 5 will refer to the more involved Birkhoff polytope $\mathcal{B}_D$ rather than the simple permutohedron. Yet, both proofs are conceptually still quite similar to the proof of Theorem 3, and can be found in the Appendix.

3.3 Target Eigenenergies

Finally, we investigate the bounds for the error ΔE_k of the individual energies. Unlike other types of errors we have seen before, ΔE_k can be negative, if $k > 0$. The reason for this also reflects the fact that there is no variational principle for individual excited states. Nonetheless, according to (17), ΔE_k is a linear function of $\tilde{\mathbf{E}}$ on the permutohedron $P(\mathbf{E})$. This means that we can use the same strategies as above to determine its lower and upper bounds.

Theorem 6. Consider the ‘Scientific Problem’ for a setting $(\mathcal{H}, \hat{H}, \mathbf{w})$ and recall (14), (17). For any k with non-degenerate w_k , the error

$\Delta E_k(\tilde{\rho}_{\mathbf{w}})$ of the k -th eigenenergy and $\Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}})$ of the ensemble energy are universally related for all $\mathbf{w}$ -ensembles $\tilde{\rho}_{\mathbf{w}}$ according to

$$c_-^{(k)}(\mathbf{w}, \mathbf{E}) \Delta E_{\mathbf{w}} \leq \Delta E_k \leq c_+^{(k)}(\mathbf{w}, \mathbf{E}) \Delta E_{\mathbf{w}}, \quad (35)$$

where

$$c_-^{(k)}(\mathbf{w}, \mathbf{E}) \equiv \begin{cases} 0, & k = 0 \\ \frac{1}{w_k - w_{k-1}}, & k > 0 \end{cases} \quad (36)$$

$$c_+^{(k)}(\mathbf{w}, \mathbf{E}) \equiv \frac{1}{w_k - w_{k+1}}. \quad (37)$$

The interpretation of these bounds is also clear. The lower bound, which is negative, is saturated when the $(k-1)$ -th and k -th energy levels mix with one another, and the positive upper bound is realized by mixing the k -th and $(k+1)$ -th levels.

The lack of a variational principle for individual excited states results in a potential error cancellation which could affect the accumulated error of all eigenenergies of interest: If one naively sums up all ΔE_k , the collective error would be grossly underestimated. To circumvent such a misleading error cancellation, one should consider instead the absolute sum $\sum_{k=0}^{D-1} |\Delta E_k|$ of various individual errors ΔE_k . By directly applying Theorem 6 D times, we obtain (while assuming a strictly decreasing $\mathbf{w}$)

$$0 \leq \sum_{k=0}^{D-1} |\Delta E_k| \leq \frac{D}{\min_{k < D-1} (w_k - w_{k+1})} \Delta E_{\mathbf{w}}. \quad (38)$$

However, much tighter bounds can be derived without invoking Theorem 6, or the polytope argument.

Theorem 7. Consider the ‘Scientific Problem’ for a setting $(\mathcal{H}, \hat{H}, \mathbf{w})$ and recall (14), (17). Then, the sum of the errors of the targeted eigenenergies and the error $\Delta E_{\mathbf{w}}(\tilde{\rho}_{\mathbf{w}})$ of the ensemble energy are universally related for all $\mathbf{w}$ -ensembles $\tilde{\rho}_{\mathbf{w}}$ in the following way. When $w_0 > w_1 > \dots > w_{D-1}$, one has

$$\begin{aligned} \frac{2\Delta E_{\mathbf{w}}}{w_0 - w_{D-1}} &\leq \sum_{k=0}^{D-1} |\Delta E_k| \\ &\leq \max_{k < D-1} \left\{ \frac{2}{w_k - w_{k+1}} \right\} \Delta E_{\mathbf{w}}. \end{aligned} \quad (39)$$

When $w_0 > w_1 > \dots > w_K = \dots = w_{D-1} = 0$ for

some $K < D-1$, one has

$$\begin{aligned} \frac{\Delta E_{\mathbf{w}}}{w_0} &\leq \sum_{k=0}^{K-1} |\Delta E_k| \\ &\leq \max_{k < K} \left\{ \frac{2\Theta(K-k-1)}{w_k - w_{k+1}} \right\} \Delta E_{\mathbf{w}}. \end{aligned} \quad (40)$$

Theorem 7 explicitly demonstrated that the absolute sum of the errors of all targeted eigenenergies is bounded from below and above by the error of the ensemble energy. Remarkably, the prefactors of various bounds involve only the weight vector $\mathbf{w}$, and are independent of the spectrum $\mathbf{E}$ of the target Hamiltonian. The proof of Theorem 7 can also be found in the Appendix.

4 Saturation of Inequalities and Optimal choice of ensemble weights

In this section we first confirm both analytically and numerically that various error bounds provided by Theorems 3 - 7 are optimal, in the sense that they can be saturated for any choice of $\mathbf{w}$, provided $\Delta E_{\mathbf{w}} \leq g_{\mathbf{w}, \mathbf{E}}$. Then, based on the explicit form $d_{\pm} \equiv d_{\pm}^{(Q)}(\mathbf{w}, \mathbf{E})$ of the prefactors of various bounds, we determine how the ensemble weights $\mathbf{w}$ need to be chosen in order to minimize the error range of the quantum states and energies targeted through the GOK variational principle. In that context, we also explain how these insights can be used in practice to improve numerical approaches in terms of the accuracy of their predictions.

4.1 Tightness of error bounds

In the following, we demonstrate by analytical means that the error bounds provided by Theorems 3 - 7 are optimal, i.e., they cannot be replaced by more restrictive bounds. We already know that such tightness is given for the bounds in Theorems 3, 6, 7 as a direct mathematical consequence of their derivation. Quite in contrast, the derivation of the bounds in Theorem 4 and 5 involved a relaxation of the optimization (13) from the underlying set $\mathcal{U}_d$ to the larger Birkhoff polytope $\mathcal{B}_d$ according to (22).

The strategy for confirming tightness is straightforward since we just need to provide for any valid $\mathbf{w}$ a corresponding $\tilde{\rho}_{\mathbf{w}}$ which saturates the corresponding inequality between the errors.

In order to illustrate this, we consider as an example the upper bound in Theorem 7 (despite the fact that its tightness is already known). Let $\mathbf{w}$ be the underlying non-degenerate weight vector and k the index for which $2/(w_l - w_{l+1})$ is maximized, i.e.,

$$k = \arg \max_{l < D-1} \left\{ \frac{2}{w_l - w_{l+1}} \right\}. \quad (41)$$

Then, according to Theorem 7, the sum of the absolute errors of various eigenenergies is bounded from above by

$$\sum_{l=0}^{D-1} |\Delta E_l| \leq \frac{2}{w_k - w_{k+1}} \Delta E_{\mathbf{w}}. \quad (42)$$

To confirm that this bound can be attained, we construct a trial state $\tilde{\rho}_{\mathbf{w}}$ for which this inequality is saturated. Our $\tilde{\rho}_{\mathbf{w}}$ is obtained from the exact state $\rho_{\mathbf{w}}$ by applying the following Jacobi rotation $\mathbf{J}_{k,k+1}(\theta)$

$$\begin{aligned} |\Psi_k\rangle &\mapsto \cos(\theta)|\Psi_k\rangle + \sin(\theta)|\Psi_{k+1}\rangle, \\ |\Psi_{k+1}\rangle &\mapsto -\sin(\theta)|\Psi_k\rangle + \cos(\theta)|\Psi_{k+1}\rangle, \end{aligned} \quad (43)$$

while keeping all other states unchanged. This leads directly to (recall Eq. (14))

$$\Delta E_{\mathbf{w}} = \sin^2(\theta)(w_k - w_{k+1})(E_{k+1} - E_k), \quad (44)$$

and

$$\sum_{l=0}^{D-1} |\Delta E_l| = 2 \sin^2(\theta)(E_{k+1} - E_k). \quad (45)$$

Comparing these two equations leads to

$$\sum_{l=0}^{D-1} |\Delta E_l| = \frac{2\Delta E_{\mathbf{w}}}{w_k - w_{k+1}}, \quad (46)$$

i.e., the absolute sum of the individual energy errors is proportional to the ensemble error, exactly by the prefactor stated in Theorem 7. Moreover, based on (44) and by recalling (26) we conclude that $\mathbf{w}$ -ensemble states $\tilde{\rho}_{\mathbf{w}}$ saturating the bound in Theorem 7 exist at least as long as $\Delta E_{\mathbf{w}} \leq (w_k - w_{k+1})(E_{k+1} - E_k)$. It is worth noticing that the right-hand side is never smaller than $g_{\mathbf{w}, \mathbf{E}}$ in (26). Last but not least, it is not difficult to see that such specific ‘saturating’ Jacobi rotations (43) also exist for various other bounds, in particular for those of Theorem 4 and 5 and also for the second case of weight vectors with

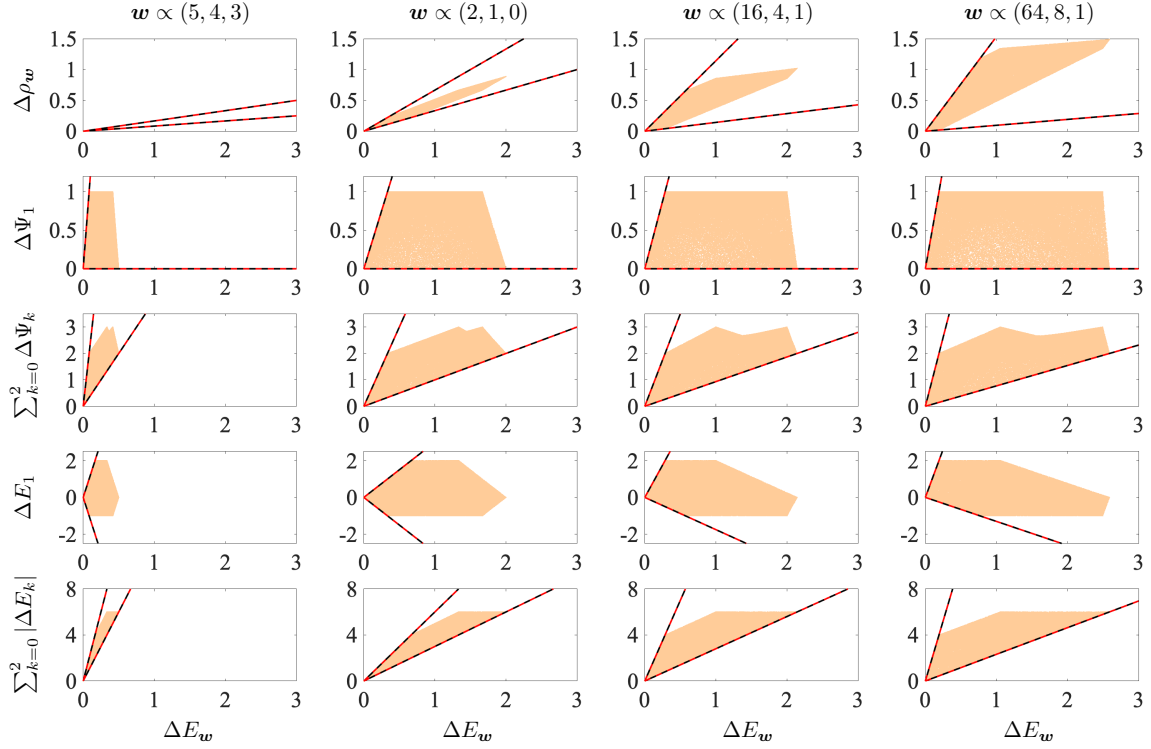


Figure 6: Numerical demonstration of the error bounds predicted by Theorems 3 - 7 and their tightness. For an energy spectrum $\mathbf{E} = (-1, 0, 2)$ and four exemplary weight vectors $\mathbf{w}$, the relations between various errors and the ensemble error ΔE_w are presented for 10^5 randomly sampled $\mathbf{w}$ -ensemble states $\tilde{\rho}_w = \hat{U}\rho_w\hat{U}^\dagger$ (orange points). Theoretical bounds (black solid lines) turn out to coincide with the actual bounds given by the sampled data (red dashed lines).

$w_0 > \dots > w_{K-1} > w_K = \dots w_{D-1} = 0$. This eventually confirms that all our error bounds are optimal in the sense that they are mathematically tight, provided the ensemble error ΔE_w is small enough, i.e., $\Delta E_w \leq g_{\mathbf{w}, \mathbf{E}}$. The numerical analysis in the subsequent section and the analytical example above reveals that for some error bounds, the range of ΔE_w for which the linear bounds are tight extends even beyond $g_{\mathbf{w}, \mathbf{E}}$. Yet, from a practical point of view it does not make sense to elaborate more on that point since the relevant regime is the one of reasonably good convergence, i.e., ΔE_w sufficiently small.

4.2 Numerical illustration of error bounds

To directly showcase the universality and the tightness of our bounds, we present in Figure 6 and 7 various errors in the quantum states and eigenenergies, of randomly sampled eigenbases encoded as unitary matrices. The unitaries are parameterized as exponentials of antisymmetric matrices, whose elements are sampled uniformly in the interval $[-\pi, \pi]$. Additionally, all possible permutation matrices are included into the

samples. We test two different scenarios: (a) a Hilbert space of dimension $D = 3$ where all three eigenstates are targeted (Figure 6), and (b) a Hilbert space of dimension $D = 5$ where the lowest three eigenstates are targeted (Figure 7). For both scenarios, we choose four exemplary weight vectors $\mathbf{w}$: One that has nearly equal entries, the one given by (52) and (53) which turns out to minimize the accumulated energy error ΔE , and those characterized by $w_i/w_{i+1} = 4$, and $w_i/w_{i+1} = 8$, respectively (excluding zeros).

As the results shown in Figures 6 and 7 clearly reveal, our analytic error bounds are satisfied and can be well saturated for both scenarios. The latter is evident by the fact that the theoretical bounds (black solid lines) and the statistical bound based on the sampled data (red dashed line) coincide. Yet, we also see that already for the relatively small dimension $D = 5$, the sampling of 10^7 trial states $\tilde{\rho}_w = \hat{U}\rho_w\hat{U}^\dagger$ is not sufficient anymore to fill out the entire area of admissible pairs $(\Delta Q(\tilde{\rho}_w), \Delta E_w(\tilde{\rho}_w))$. This is due to the increase of the dimension of the manifold of unitaries as D^2 , which in turn highlights the

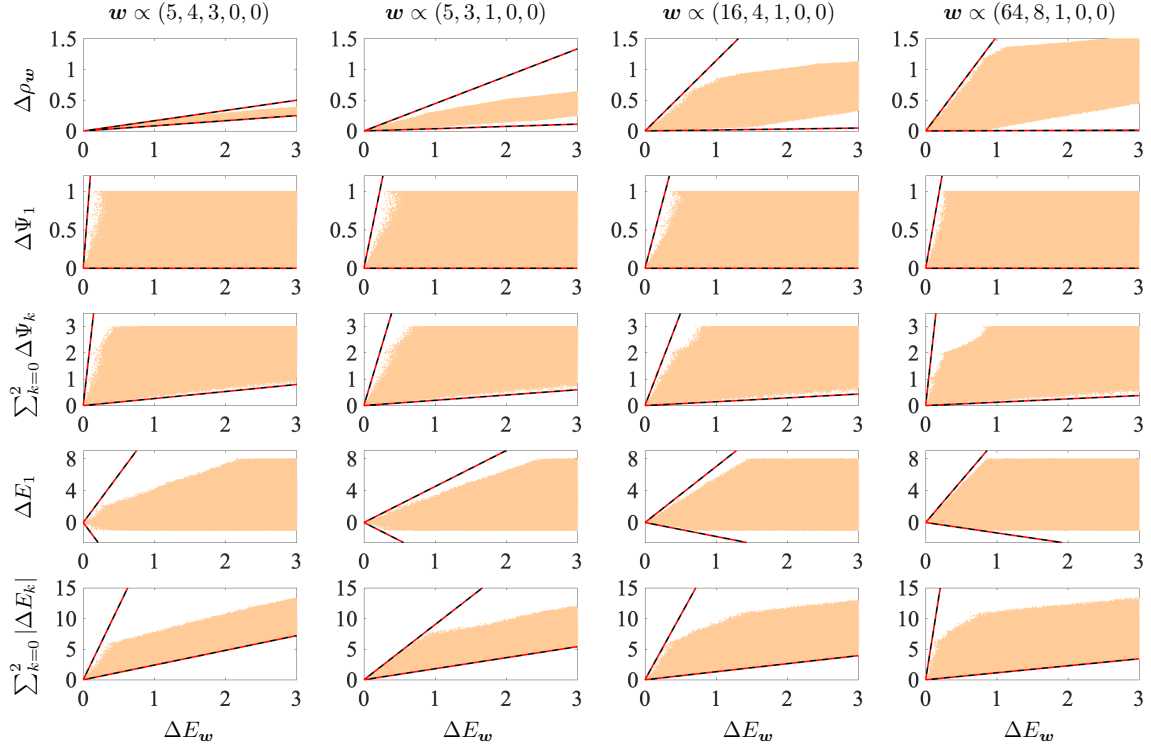


Figure 7: Numerical demonstration of the error bounds predicted by Theorems 3 - 7 and their tightness. For an energy spectrum $\mathbf{E} = (-1, 0, 2, 5, 8)$ and four exemplary weight vectors $\mathbf{w}$, the relations between various errors and the ensemble error ΔE_w are presented for 10^7 randomly sampled $\mathbf{w}$ -ensemble states $\tilde{\rho}_w = \hat{U}\rho_w\hat{U}^\dagger$ (orange points). Theoretical bounds (black solid lines) turn out to coincide with the actual bounds given by the sampled data (red dashed lines).

value of our analytical analysis in the previous and subsequent section. Moreover, by comparing the numerical results for the four exemplary weight vectors $\mathbf{w}$, we observe the following trends. First, the range of possible errors $\Delta\rho_w, \Delta\Psi, \Delta E$ defined by the respective lower and upper bounds is, the narrower the more equal various weights w_l are. In contrast, for $\Delta\Psi_k, \Delta E_k$ the range gets narrower as the weight gaps $|w_k - w_{k\pm 1}|$ get bigger. In practical applications, however, the ideal selection criteria for $\mathbf{w}$ would not refer to the narrowness of the error range but instead to the ‘flatness’ of the upper bound, i.e., the value of the corresponding prefactor. Indeed, the smaller the upper bound is, the smaller must be the error of the quantity of interest and thus the more accurate are the predictions made by the numerical approach. In that regard, for all quantities except the ensemble state, the $\mathbf{w}$ of the second column, namely an equally spaced $\mathbf{w}$, seems to be superior. Since the latter observation might be subject to the specific energy spectra chosen here, we will elaborate in the next section more on the crucial point of weight selection. In partic-

ular, we will succeed in determining by analytical means the optimal weight vectors for each quantity of interest, and their potential dependence on the underlying energy spectrum $\mathbf{E}$.

4.3 Optimal choice of weights

The main results of our work, the explicit expressions of the linear error bounds in Theorems 3 - 7 allow us to determine in the following the optimal auxiliary weights w_k for practical applications. Since we know from the theoretical and numerical analyses in the previous sections that various error bounds are tight, we just need to minimize for each quantity Q of interest the prefactor $d_+^{(Q)}(\mathbf{w}, \mathbf{E})$ of its upper bound. In that way, we will identify the optimal vectors $\mathbf{w}$ for which the predictive power of any numerical approach based on the GOK variational principle is maximized: given an error ΔE_w of the ensemble energy, these are exactly those $\mathbf{w}$ for which the largest possible error ΔQ that Q may have is minimal.

Before we derive the optimal weight vectors, two crucial comments are in order here concern-

Quantity	Error	Theorem	Lowest Upper Bound	Optimal $\mathbf{w}$
$ \Psi_k\rangle$	$\Delta\Psi_k$	4	$k(r_- + r_+) + r_+$	$\mathbf{w} \propto (\underbrace{r_- + r_+, \dots, r_- + r_+}_{k-1}, \underbrace{r_+, 0, \dots, 0}_{D-k})$
$\{ \Psi_k\rangle\}_{k=0}^{D-1}$	$\sum_{k=0}^{D-1} \Delta\Psi_k$	5	$\sum_{k=1}^{D-1} 2k(E_k - E_{k-1})^{-1}$	$\mathbf{w} \propto (\sum_{l=k+1}^{D-1} (E_l - E_{l-1})^{-1})_{k=0}^{D-1}$
$\{ \Psi_k\rangle\}_{k=0}^{K-1}$	$\sum_{k=0}^{K-1} \Delta\Psi_k$	5	$\sum_{k=1}^K 2k(E_k - E_{k-1})^{-1} \Theta(K-k)$	$\mathbf{w} \propto (\sum_{l=k+1}^{D-1} (E_l - E_{l-1})^{-1} \Theta(K-l))_{k=0}^{D-1}$
E_k	$ \Delta E_k $	6	$2k + 1$	$\mathbf{w} \propto (\underbrace{2, \dots, 2}_k, \underbrace{1, 0, \dots, 0}_{D-k-1})$
$\{E_k\}_{k=0}^{D-1}$	$\sum_{k=0}^{D-1} \Delta E_k $	7	$D(D-1)$	$\mathbf{w} \propto (D-1, D-2, \dots, 1, 0)$
$\{E_k\}_{k=0}^{K-1}$	$\sum_{k=0}^{K-1} \Delta E_k $	7	K^2	$\mathbf{w} \propto (2K-1, 2K-3, \dots, 1, 0, \dots, 0)$

Table 1: The lowest upper bounds and the corresponding optimal choice of $\mathbf{w}$ (normalized to 1) for various errors of the quantum states and eigenenergies. We defined $r_{\pm} \equiv |E_k - E_{k\pm 1}|^{-1}$, $\Theta(x)$ denotes the Heaviside step function (32), and K satisfies $K < D-1$.

ing the practical relevance of the anticipated results and our error bounds in general. At first sight, one may wonder how our results can be applied in practice, given that they necessitate the approximate knowledge of $\Delta E_{\mathbf{w}}$ and for some quantities Q even $\mathbf{E}$. First, as far as $\Delta E_{\mathbf{w}}$ is concerned, we recall that in numerous variational optimization methods systematic extrapolation techniques are available for determining good estimates of the exact variational energies. Exactly the same schemes could be used in the context of the GOK variational principle. Prime examples would be the schemes in the density matrix renormalization group (DMRG) ansatz for extrapolating to the results of infinite bond-dimension [75–82] and in the selected configuration interaction approach for extrapolating to the full configuration interaction ansatz involving all electron configurations [83, 84]. Second, for those quantities Q whose optimal $\mathbf{w}$ will depend on the energy spectrum $\mathbf{E}$, one could get a reasonable estimate of $\mathbf{E}$ through a low-level method such as the Hartree-Fock ansatz or any cheaper post Hartree-Fock approach. Most importantly, this initial estimate of $\mathbf{E}$ and resulting $\mathbf{w}$ can further be improved during the course of the variational minimization of the ensemble energy. We explain this exemplarily for the state-average CASSCF approach [50–52, 54]. There, one first optimizes the active orbitals via a gradient step in order to minimize the ensemble energy $E_{\mathbf{w}}$. Then, in a second step, one determines the targeted eigenstates and their energies within the active space of the updated orbitals through some exact or ap-

proximate diagonalization technique. This two-step procedure is repeated numerous times until convergence is reached. Apparently, one obtains more and more accurate estimates of the energy spectrum $\mathbf{E}$ during the optimization and could therefore constantly update $\mathbf{w}$ in order to maximize the predictive power of the final results.

In the following we present and discuss the optimal weight vectors which reduce the error ranges of $\Delta E_k, \Delta E, \Delta\Psi_k$ and $\Delta\Psi$, respectively. We remark that the underlying reasoning, however, would not make any sense for the error $\Delta\rho_{\mathbf{w}}$ of the ensemble state, since $\rho_{\mathbf{w}}$ itself explicitly depends on $\mathbf{w}$. In particular the upper error bound in Theorem 3 vanishes for $\mathbf{w} \propto (1, 1, \dots, 1)$ because there exists only one quantum state $\tilde{\rho}_{\mathbf{w}} = \rho_{\mathbf{w}} = \mathbb{1}/D$ for that specific weight vector. We commence with the practically most important case of targeting individual energies E_k . Since the error ΔE_k can be also negative, we first merge both bounds in Theorem 6 to a single one,

$$0 \leq |\Delta E_k| \leq \max \{ \mu_{k-1}^{-1}, \mu_k^{-1} \} \Delta E_{\mathbf{w}}. \quad (47)$$

Here, we introduced the more instructive parameters

$$\begin{aligned} \mu_l &\equiv w_l - w_{l+1}, \quad l < D-1 \\ \mu_{D-1} &\equiv w_{D-1}, \end{aligned} \quad (48)$$

where obviously $\mu_l \geq 0$ for all l since $\mathbf{w} = \mathbf{w}^\downarrow$. Additionally, the normalization condition translates to

$$\sum_{l=0}^{D-1} w_l = \sum_{l=0}^{D-1} (l+1) \mu_l = 1, \quad (49)$$

which implies in particular $\mu_l \leq 1/(l+1)$. Throughout this section, any μ is assumed to respect these conditions such that the corresponding $\mathbf{w}$ is indeed a probability distribution. Minimizing then the prefactor of the right-hand side in (47) over all admissible μ yields in a straightforward manner

$$\min_{\mu: \mu_{k-1}, \mu_k > 0} \max \left\{ \mu_{k-1}^{-1}, \mu_k^{-1} \right\} = 2k+1, \quad (50)$$

realized by $\mu_k = \mu_{k+1} = (2k+1)^{-1}$. This translates to the optimal weight vector given by

$$\mathbf{w} = \frac{1}{2k+1} \underbrace{(2, \dots, 2)}_k, \underbrace{1, 0, \dots, 0)}_{D-k-1}. \quad (51)$$

In this vector, the gaps between those consecutive weights which involve the target eigenstate (k) are maximized, while all other gaps are minimized to 0. The corresponding lowest upper bound for $\Delta E_k / \Delta E_{\mathbf{w}}$ is $2k+1$, which increases linearly in k . Remarkably, this rationalizes in quantitative terms the common expectation that the higher an excitation lies, the larger can be the error of the corresponding energy E_k .

In the following we present and discuss the optimal $\mathbf{w}$ and the corresponding lowest upper error bounds also for the other relevant quantities. These additional key results of our work can be derived similarly to those for E_k . Due to the technical character of their derivations we refer the reader to Appendix B for various details.

If we are interested in calculating all D eigenenergies accurately, the optimal $\mathbf{w}$ is determined by minimizing the upper bound of $\sum_{k=0}^{D-1} |\Delta E_k|$ in Theorem 7. By recalling the optimal $\mathbf{w}$ for an individual energy E_k in (51), it is less surprisingly that the optimal $\mathbf{w}$ follows here as

$$\mathbf{w} = \frac{2}{D(D-1)} (D-1, D-2, \dots, 1, 0). \quad (52)$$

This means that all gaps between consecutive weights are equal and thus equally maximal. The corresponding lowest upper bound in Theorem 7 follows as $D(D-1)$. If we are interested in calculating just the lowest K ($0 < K < D-1$) eigenenergies, which is the more sensible setting for realistic quantum systems with exceedingly large D , the following $\mathbf{w}$ turns out to be optimal

$$\mathbf{w} = \frac{1}{K^2} (2K-1, 2K-3, \dots, 3, 1, 0, \dots, 0). \quad (53)$$

This result makes sense as well from an intuitive point of view: for the lowest K energies, the weights between consecutive energies are maximally distinct which is achieved also by assigning the weight 0 to all higher energy levels. The corresponding lowest upper bound for the ratio $\sum_{k=0}^{K-1} |\Delta E_k| / \Delta E_{\mathbf{w}}$ follows directly from (53) and Theorem 7 as K^2 , which grows quadratically as function of K .

To discuss the case of the k -th eigenstate $|\Psi_k\rangle$, we recast the upper bound in Theorem 4 as

$$\Delta \Psi_k \leq \max \left\{ \frac{r_-}{\mu_{K-1}}, \frac{r_+}{\mu_k} \right\}, \quad (54)$$

where $r_{\pm} \equiv |E_k - E_{k\pm 1}|^{-1}$. The lowest upper bound then follows as

$$\min_{\mu: \mu_{K-1}, \mu_K > 0} \max \left\{ \frac{r_-}{\mu_{K-1}}, \frac{r_+}{\mu_k} \right\} = k(r_- + r_+) + r_+, \quad (55)$$

with $\mu_{K-1} = r_- / [k(r_- + r_+) + r_+]$ and $\mu_k = r_+ / [k(r_- + r_+) + r_+]$. This leads to the optimal weight vector given by

$$\mathbf{w} = \frac{1}{k(r_- + r_+) + r_+} \times \underbrace{(r_- + r_+, \dots, r_- + r_+)}_k, \underbrace{r_+, 0, \dots, 0)}_{D-k-1}. \quad (56)$$

In contrast to the optimal weights for ΔE_k and ΔE , which are independent of the energy spectrum $\mathbf{E}$, the optimal gaps μ_j between consecutive weights are now balanced by the corresponding energy gaps. It are not the gaps that are maximized by the optimal weight vector but the product of the weight and energy gaps. In a similar fashion we find also the optimal weights for targeting multiple eigenstates. When all D eigenstates are considered, the optimal $\mathbf{w}$ for achieving the lowest upper bound for $\sum_{k=0}^{D-1} \Delta \Psi_k$ follows as

$$\mathbf{w} \propto \left(\sum_{j=k+1}^{D-1} (E_j - E_{j-1})^{-1} \right)_{k=0}^{D-1}, \quad (57)$$

whereas when only the lowest $K < D-1$ eigenstates are considered, the optimal $\mathbf{w}$ for achieving the lowest upper bound for $\sum_{k=0}^{K-1} \Delta \Psi_k$ follows as

$$\mathbf{w} \propto \left(\sum_{j=k+1}^{D-1} (E_j - E_{j-1})^{-1} \Theta(K-j) \right)_{k=0}^{D-1}. \quad (58)$$

We summarize the optimal weight vectors and lowest upper bounds for all relevant quantities in Table 1. It is worth recalling here that this second key accomplishment of our work was only possible as we succeeded in the first place in deriving the universal and optimal error bounds presented in Theorems 3-7. With the ‘instruction manual’ Table 1 at hand, the optimal $\mathbf{w}$ can be chosen for any quantity Q of interest. This will then maximize the accuracy of any numerical approach based on the GOK variational principle such as GOK-DFT [25–37], $\mathbf{w}$ -RDMFT [42–46], specific quantum Monte Carlo techniques [47–49], traditional quantum chemical methods [50–54] and quantum computing [55–63].

5 Illustration: Variational Quantum Eigensolver

To showcase the practical relevance of our results, we apply our theorems within the framework of variational quantum eigensolvers (VQE) [85], used to solve the transverse Ising model

$$\hat{H} = \sum_{i=1}^N a_i \hat{X}_i + \sum_{1 \leq i < j \leq N} J_{ij} \hat{Z}_i \hat{Z}_j, \quad (59)$$

where $\hat{X}_i, \hat{Z}_i$ denote the Pauli operators for the i -th spin. For this proof of concept, it is sufficient to choose $N = 2$, resulting in a Hilbert space of dimension $D = 2^2$. The coefficients a_i and J_{ij} are fixed random numbers uniformly sampled from the interval $[0, 1)$. In our work they are given by $J_{12} = 0.09000, a_1 = 0.32696, a_2 = 0.80430$ which results in the following eigenenergies

$$\begin{aligned} E_0 &= -1.13483, \\ E_1 &= -0.48575, \\ E_2 &= 0.48575, \\ E_3 &= 1.13483. \end{aligned} \quad (60)$$

The quantum simulations were conducted using PennyLane [86], a quantum machine learning and simulation library. To diagonalize the Hamiltonian (59) in virtue of the ensemble variational principle through a unitary transformation $\hat{U}$, we resort to the parametric ansatz proposed in Ref. [87]. By increasing the circuit depth, the ansatz for $\hat{U}$ is then capable of diagonalizing the Hamiltonian (59) with arbitrary precision. Three different weight vectors $\mathbf{w}$ are employed,

$$\mathbf{w}^{(n)} \propto (4^n, 3^n, 2^n, 1^n), \quad n = 1, 2, 3 \quad (61)$$

to demonstrate the effect of the choice of $\mathbf{w}$ in the error of the quantum states and eigenenergies during the optimization of the circuit parameters using the Adam algorithm [88].

In Figure 8, we present the errors of the ensemble state ($\Delta\rho_{\mathbf{w}}$), the first excited state ($\Delta\Psi_1$) and eigenenergy (ΔE_1), as well as the accumulated errors of various eigenstates ($\Delta\Psi \equiv \sum_{k=0}^3 \Delta\Psi_k$) and eigenenergies ($\Delta E \equiv \sum_{k=0}^3 |\Delta E_k|$), against the error $\Delta E_{\mathbf{w}}$ of the ensemble energy. The red data points record the evolution of the errors during the optimization of the circuit parameters, while the upper and lower bounds of the errors given by Theorems 3 - 7 are shown as black lines. In contrast to Figure 6 and 7, we opted here for a double-logarithmic scale for various error types except for ΔE_1 (which can be negative), in order to resolve the behavior of the errors in the important regime where $\Delta E_{\mathbf{w}}$ is arbitrarily small (regime of convergence).

By following in each plot from the right side the red data points we observe that during the course of the optimization the variational ensemble energy $E_{\mathbf{w}}$ gets more and more accurate. Since the individual and accumulated errors ΔQ of the eigenenergies and eigenstates are linearly bounded from above by $\Delta E_{\mathbf{w}}$ according to Theorems 3 - 7, the excellent convergence of the latter enforces an equally excellent (or even better) convergence of the former. In particular, for the error quantities which are bounded from above and below by positive bounds (first, third and fifth row), the corresponding black lines form “channels” which “guides” the respective errors ΔQ to the regime of excellence convergence. Note here, that the linear positive bounds $d_- \Delta E_{\mathbf{w}} \leq \Delta Q \leq d_+ \Delta E_{\mathbf{w}}$ in Theorems 3, 5, 7 translate here to

$$\begin{aligned} \log(d_-) + \log(\Delta E_{\mathbf{w}}) &\leq \log(\Delta Q) \\ &\leq \log(d_+) + \log(\Delta E_{\mathbf{w}}). \end{aligned} \quad (62)$$

Accordingly, the vertical arrangement of any black line is defined by an offset $\log(d_{\pm})$ and the width of each “channel” follows as $\log(d_+) - \log(d_-)$. The other two error quantities (second and fourth row) are also bounded from above by the ensemble error but they do not have a positive lower bound. This means that they may vanish even if the ensemble energy has an error $\Delta E_{\mathbf{w}} > 0$. Indeed, this occasionally happens during the course of the optimization for the error

ΔE_1 which changes signs from time to time.

Comparison of the three columns confirms again that the choice of the weights can have a strong impact on the narrowness of the error “channels” of $\Delta\rho_w, \Delta\Psi, \Delta E$ and the vertical arrangements of various black lines. For instance, according to the bounds in Theorem 3, the “channel” of $\Delta\rho_w$ could become arbitrarily narrow by making various weights more and more equal. Even changing the weights from those of the last two columns to just an equally spaced $\mathbf{w}$ (first column) has already a huge effect on the narrowness. Yet, in that context one must not forget that changing $\mathbf{w}$ even further to a constant weight vector $\mathbf{w} \propto (1, \dots, 1)$ would trivialise the search space underlying the GOK variational principle (as explained above) and thus the latter would lose its predictive power entirely. Moreover, we observe in agreement with the predictions made by Table 1 that the first weight vector, which is close to the optimal one for the accumulated energy error ΔE , lowers indeed the upper bound compared to the other two $\mathbf{w}$, yet the effect is not that drastic. At the same time, this equally spaced $\mathbf{w}$ is far from being the optimal choice for the other error quantities (c.f Table 1). For instance, in the fourth row a wider error range follows compared to the second and third column. That is not surprising since the tightest bounds on ΔE_1 are those obtained by maximizing either $w_1 - w_2$ (upper bound) or $w_0 - w_1$ (lower bound) with the effect of sacrificing the accuracy of all other eigenstates and eigenenergies. This emphasizes the conceptual difference between the errors of individual excited states and the error of the ensemble state. This in turn suggests a potential for error cancellation due to the absence of a variational principle for individual excited states. All these effects will be even more well-pronounced in bigger systems with larger Hilbert spaces.

Finally, we observe that many of the error bounds in our example are occasionally (approximately) saturated during the optimization of the circuit parameters. Such a saturation actually occurs for all three weight vectors. This further demonstrates the particular relevance of our analytic error bounds which apparently can dictate the behavior of the errors ΔQ as function of ΔE_w . The worst-case scenario where the error of the eigenstates or eigenenergies is maximized for a fixed ΔE_w could indeed occur during the op-

timization process. From this concrete example, we therefore conclude that the weight vector $\mathbf{w}$ of the ensemble state is more than an auxiliary quantity, and that a carefully chosen $\mathbf{w}$ can lead to drastic improvements in the errors of the individual eigenstates and eigenenergies, and significantly improve the predictive power of the GOK variational principle.

6 Summary and Conclusions

We have provided a sound basis for the ensemble variational principle [21] which has been proposed by Gross, Oliveira and Kohn as a seminal extension of the famous Rayleigh-Ritz variational principle. It minimizes the energy expectation value $\text{Tr}[\tilde{\rho}_w \hat{H}]$ of a quantum system over ensemble states $\tilde{\rho}_w \equiv \sum_{k=0}^{K-1} w_k |\tilde{\Psi}_k\rangle\langle\tilde{\Psi}_k|$ with fixed auxiliary spectrum $\mathbf{w}$ in order to target excited states by variational means. By employing elegant concepts from convex geometry we confirmed by constructive means the predictive power of this GOK variational principle and proved the validity of the underlying critical hypothesis: Whenever the ensemble energy is well-converged, the same holds true for the individual eigenstates and eigenenergies E_k , provided various auxiliary weights w_k were chosen differently. To be more specific, as our main accomplishment, we derived linear bounds

$$d_-^{(Q)}(\mathbf{w}, \mathbf{E}) \Delta E_w \leq \Delta Q \leq d_+^{(Q)}(\mathbf{w}, \mathbf{E}) \Delta E_w \quad (63)$$

on the errors ΔQ of these sought-after quantities. A detailed analytical and numerical analysis in Section 4 has confirmed the tightness and practical significance of these error bounds. For instance, in practical applications such as the variational quantum eigensolver (Section 5), they dictate the convergence rate of various physical quantities as function of the error of the variational ensemble energy E_w .

Moreover, a comprehensive analysis of the explicit form of the prefactors $d_{\pm}^{(Q)}(\mathbf{w}, \mathbf{E})$ allowed us for each physical quantity of interest to determine the corresponding optimal auxiliary weights $\mathbf{w}$. The corresponding Table 1 which summarizes these key results will serve as an ‘instruction manual’ for practical applications. Following it will then maximize the predictive power of the GOK variational principle, and thus the accuracy of any ensemble-based numerical approach, such as

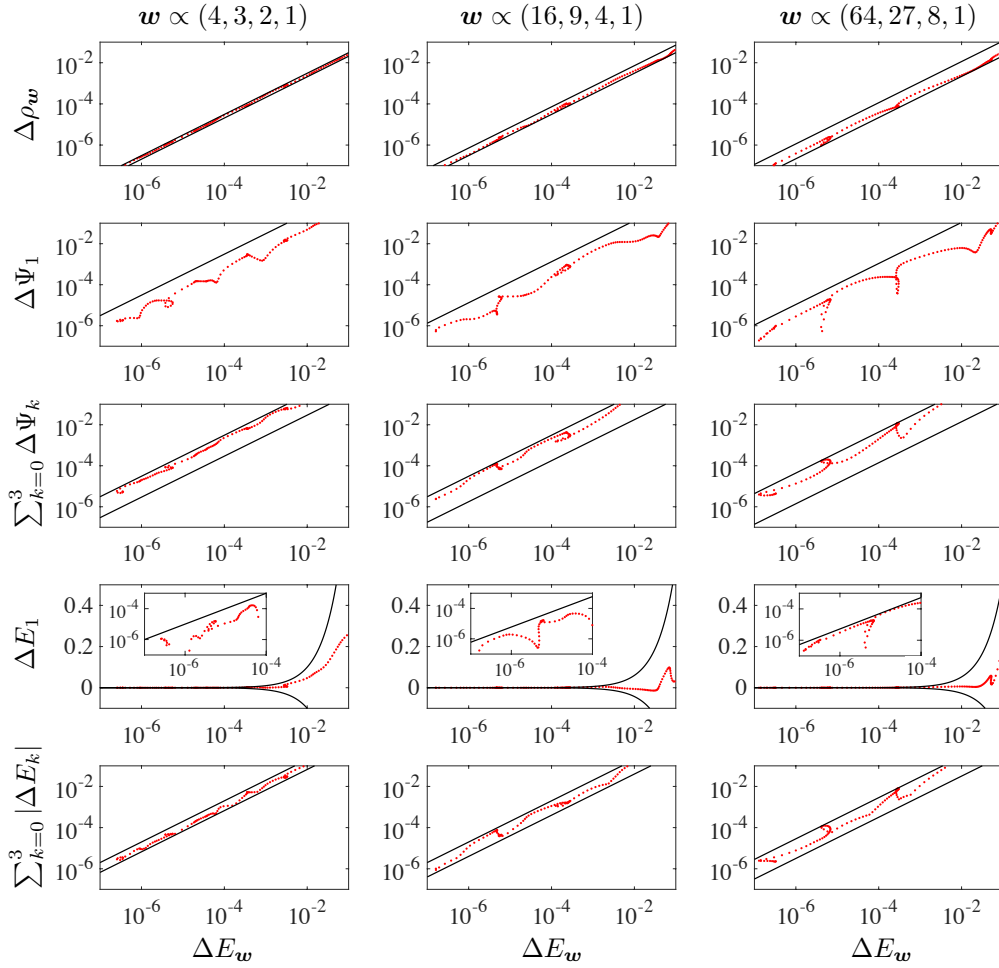


Figure 8: Illustration of the practical relevance of our main results, Theorems 3 - 7, in the context of a variational quantum eigensolver applied to the transverse field Ising model: During the course of optimization (results depicted as red points), our theoretical bounds shown as black solid lines guarantee convergence of various relevant quantities as the ensemble error ΔE_w gets arbitrarily small.

GOK-DFT [25–37], w -RDMFT [42–46], specific quantum Monte Carlo techniques [47–49], traditional quantum chemical methods [50–54] and quantum computing [55–63].

In summary, our work cemented the GOK variational principle as the cornerstone for excited state methods in complete analogy to the Rayleigh-Ritz variational principle for ground state methods. From a practical point of view, our analytical error bounds reveal the fundamental guiding principle for designing optimal weight vectors w for computing excited states and energies in physics, chemistry and materials science.

Acknowledgments

We are grateful to C.L. Benavides-Riveros, K. Liao, J. Liebert, S. Paeckel and Z. Xie for inspiring discussions. We acknowledge financial

support from the German Research Foundation (Grant SCHI 1476/1-1), the Munich Center for Quantum Science and Technology, and the Munich Quantum Valley, which is supported by the Bavarian state government with funds from the Hightech Agenda Bayern Plus.

References

- [1] M. Green, A. Ho-Baillie, and H. Snaith. The emergence of perovskite solar cells. *Nature Photon.*, 8(7):506, 2014. DOI: [10.1038/nphoton.2014.134](https://doi.org/10.1038/nphoton.2014.134).
- [2] P. J. M. Johnson, A. Halpin, T. Morizumi, V. I. Prokhorenko, O. P. Ernst, and R. J. D. Miller. Local vibrational coherences drive the primary photochemistry of vision. *Nat. Chem.*, 7(12):980, 2015. DOI: [10.1038/nchem.2398](https://doi.org/10.1038/nchem.2398).

- [3] V. W.-W. Yam. Using synthesis to steer excited states and their properties and functions. *Nat. Synth.*, 2(2):94, 2023. DOI: [10.1038/s44160-022-00202-5](https://doi.org/10.1038/s44160-022-00202-5).
- [4] Y.-C. Cheng and G. R. Fleming. Dynamics of light harvesting in photosynthesis. *Annu. Rev. Phys. Chem.*, 60(1):241, 2009. DOI: [10.1146/annurev.physchem.040808.090259](https://doi.org/10.1146/annurev.physchem.040808.090259).
- [5] G. Cerullo, D. Polli, G. Lanzani, S. De Silvestri, H. Hashimoto, and R. J. Cogdell. Photosynthetic Light Harvesting by Carotenoids: Detection of an Intermediate Excited State. *Science*, 298(5602):2395, 2002. DOI: [10.1126/science.1074685](https://doi.org/10.1126/science.1074685).
- [6] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull. Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms. *Phys. Rev. X*, 5:041041, 2015. DOI: [10.1103/PhysRevX.5.041041](https://doi.org/10.1103/PhysRevX.5.041041).
- [7] M. Motta, C. Genovese, F. Ma, Z.-H. Cui, R. Sawaya, G. K.-L. Chan, N. Chepiga, P. Helms, C. Jiménez-Hoyos, A. J. Millis, U. Ray, E. Ronca, H. Shi, S. Sorella, E. M. Stoudenmire, S. R. White, and S. Zhang. Ground-State Properties of the Hydrogen Chain: Dimerization, Insulator-to-Metal Transition, and Magnetic Phases. *Phys. Rev. X*, 10:031058, 2020. DOI: [10.1103/PhysRevX.10.031058](https://doi.org/10.1103/PhysRevX.10.031058).
- [8] G. Onida, L. Reining, and A. Rubio. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.*, 74:601, 2002. DOI: [10.1103/RevModPhys.74.601](https://doi.org/10.1103/RevModPhys.74.601).
- [9] L. Serrano-Andrés and M. Merchán. Quantum chemistry of the excited state: 2005 overview. *J. Mol. Struct.: THEOCHEM*, 729(1):99, 2005. DOI: [10.1016/j.theochem.2005.03.020](https://doi.org/10.1016/j.theochem.2005.03.020).
- [10] H. Lischka, D. Nachtigallová, Adélia J. A. Aquino, P. G. Szalay, F. Plasser, F. B. C. Machado, and M. Barbatti. Multireference Approaches for Excited States of Molecules. *Chem. Rev.*, 118(15):7293, 2018. DOI: [10.1021/acs.chemrev.8b00244](https://doi.org/10.1021/acs.chemrev.8b00244).
- [11] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi. Excited States with Selected Configuration Interaction-Quantum Monte Carlo: Chemically Accurate Excitation Energies and Geometries. *J. Chem. Theory Comput.*, 15(9):4896, 2019. DOI: [10.1021/acs.jctc.9b00476](https://doi.org/10.1021/acs.jctc.9b00476).
- [12] L. Pausch, E. G. Carnio, A. Rodríguez, and A. Buchleitner. Chaos and Ergodicity across the Energy Spectrum of Interacting Bosons. *Phys. Rev. Lett.*, 126:150601, 2021. DOI: [10.1103/PhysRevLett.126.150601](https://doi.org/10.1103/PhysRevLett.126.150601).
- [13] A. Baiardi, A. K. Kelemen, and M. Reiher. Excited-State DMRG Made Simple with FEAST. *J. Chem. Theory Comput.*, 18(1):415, 2022. DOI: [10.1021/acs.jctc.1c00984](https://doi.org/10.1021/acs.jctc.1c00984).
- [14] J. Westermayr and P. Marquetand. Machine learning for electronically excited states of molecules. *Chemical Reviews*, 121(16):9873, 2021. DOI: [10.1021/acs.chemrev.0c00749](https://doi.org/10.1021/acs.chemrev.0c00749).
- [15] P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin. A mountaineering strategy to excited states: Highly accurate reference energies and benchmarks. *J. Chem. Theory Comput.*, 14(8):4360–4379, 2018. DOI: [10.1021/acs.jctc.8b00406](https://doi.org/10.1021/acs.jctc.8b00406).
- [16] L. González, D. Escudero, and L. Serrano-Andrés. Progress and Challenges in the Calculation of Electronic Excited States. *ChemPhysChem*, 13(1):28, 2012. DOI: [10.1002/cphc.201100200](https://doi.org/10.1002/cphc.201100200).
- [17] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136:B864, Nov 1964. DOI: [10.1103/PhysRev.136.B864](https://doi.org/10.1103/PhysRev.136.B864).
- [18] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77:259, 2005. DOI: [10.1103/RevModPhys.77.259](https://doi.org/10.1103/RevModPhys.77.259).
- [19] G. Carleo and M. Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602, 2017. DOI: [10.1126/science.aag2302](https://doi.org/10.1126/science.aag2302).
- [20] C. L. Benavides-Riveros, N. N. Lathiotakis, C. Schilling, and M. A. L. Marques. Relating correlation measures: The importance of the energy gap. *Phys. Rev. A*, 95:032507, 2017. DOI: [10.1103/PhysRevA.95.032507](https://doi.org/10.1103/PhysRevA.95.032507).
- [21] E. K. U. Gross, L. N. Oliveira, and W. Kohn. Rayleigh-Ritz variational principle for en-

- sembles of fractionally occupied states. *Phys. Rev. A*, 37:2805, 1988. DOI: [10.1103/PhysRevA.37.2805](https://doi.org/10.1103/PhysRevA.37.2805).
- [22] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864, 1964. DOI: [10.1103/PhysRev.136.B864](https://doi.org/10.1103/PhysRev.136.B864).
- [23] E. Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52:997–1000, 1984. DOI: [10.1103/PhysRevLett.52.997](https://doi.org/10.1103/PhysRevLett.52.997).
- [24] M. Petersilka, U. J. Gossmann, and E. K. U. Gross. Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.*, 76:1212–1215, 1996. DOI: [10.1103/PhysRevLett.76.1212](https://doi.org/10.1103/PhysRevLett.76.1212).
- [25] E. K. U. Gross, L. N. Oliveira, and W. Kohn. Density-functional theory for ensembles of fractionally occupied states. I. Basic formalism. *Phys. Rev. A*, 37:2809, 1988. DOI: [10.1103/PhysRevA.37.2809](https://doi.org/10.1103/PhysRevA.37.2809).
- [26] Z. Yang, A. Pribram-Jones, K. Burke, and C. A. Ullrich. Direct Extraction of Excitation Energies from Ensemble Density-Functional Theory. *Phys. Rev. Lett.*, 119:033003, 2017. DOI: [10.1103/PhysRevLett.119.033003](https://doi.org/10.1103/PhysRevLett.119.033003).
- [27] K. Deur, L. Mazouin, and E. Fromager. Exact ensemble density functional theory for excited states in a model system: Investigating the weight dependence of the correlation energy. *Phys. Rev. B*, 95:035120, 2017. DOI: [10.1103/PhysRevB.95.035120](https://doi.org/10.1103/PhysRevB.95.035120).
- [28] B. Senjean and E. Fromager. Unified formulation of fundamental and optical gap problems in density-functional theory for ensembles. *Phys. Rev. A*, 98:022513, 2018. DOI: [10.1103/PhysRevA.98.022513](https://doi.org/10.1103/PhysRevA.98.022513).
- [29] P.-F. Loos and E. Fromager. A weight-dependent local correlation density-functional approximation for ensembles. *J. Chem. Phys.*, 152(21), 2020. DOI: [10.1063/5.0007388](https://doi.org/10.1063/5.0007388).
- [30] B. Senjean and E. Fromager. N-centered ensemble density-functional theory for open systems. *Int. J. Quantum Chem.*, 120(21):e26190, 2020. DOI: [10.1002/qua.26190](https://doi.org/10.1002/qua.26190).
- [31] M. J. P. Hodgson, J. Wetherell, and Emmanuel Fromager. Exact exchange-correlation potentials for calculating the fundamental gap with a fixed number of electrons. *Phys. Rev. A*, 103:012806, 2021. DOI: [10.1103/PhysRevA.103.012806](https://doi.org/10.1103/PhysRevA.103.012806).
- [32] E. Kraisler, M. J. P. Hodgson, and E. K. U. Gross. From Kohn–Sham to many-electron energies via step structures in the exchange-correlation potential. *J. Chem. Theory Comput.*, 17(3):1390–1407, 2021. DOI: [10.1021/acs.jctc.0c01093](https://doi.org/10.1021/acs.jctc.0c01093).
- [33] T. Gould and L. Kronik. Ensemble generalized Kohn–Sham theory: The good, the bad, and the ugly. *J. Chem. Phys.*, 154(9), 2021. DOI: [10.1063/5.0040447](https://doi.org/10.1063/5.0040447).
- [34] Z.-H. Yang. Second-order perturbative correlation energy functional in the ensemble density-functional theory. *Phys. Rev. A*, 104:052806, 2021. DOI: [10.1103/PhysRevA.104.052806](https://doi.org/10.1103/PhysRevA.104.052806).
- [35] Y. Lu and J. Gao. Multistate density functional theory of excited states. *J. Phys. Chem. Lett.*, 13(33):7762–7769, 2022. DOI: [10.1021/acs.jpcllett.2c02088](https://doi.org/10.1021/acs.jpcllett.2c02088).
- [36] S. Giarrusso and P.-F. Loos. Exact excited-state functionals of the asymmetric Hubbard dimer. *J. Phys. Chem. Lett.*, 14(39):8780–8786, 2023. DOI: [10.1021/acs.jpcllett.3c02052](https://doi.org/10.1021/acs.jpcllett.3c02052).
- [37] F. Cernatic, P.-F. Loos, B. Senjean, and E. Fromager. Neutral electronic excitations and derivative discontinuities: An extended n -centered ensemble density functional theory perspective. *Phys. Rev. B*, 109:235113, Jun 2024. DOI: [10.1103/PhysRevB.109.235113](https://doi.org/10.1103/PhysRevB.109.235113).
- [38] T. Gilbert. Hohenberg-Kohn theorem for nonlocal external potentials. *Phys. Rev. B*, 12:2111, 1975. DOI: [10.1103/PhysRevB.12.2111](https://doi.org/10.1103/PhysRevB.12.2111).
- [39] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v -representability problem. *Proc. Natl. Acad. Sci. U.S.A.*, 76(12):6062, 1979. DOI: [10.1073/pnas.76.12.6062](https://doi.org/10.1073/pnas.76.12.6062).
- [40] S. M. Valone. Consequences of extending 1-matrix energy functionals from pure-state representable to all ensemble representable 1-matrices. *J. Chem. Phys.*, 73(3):1344, 1980. DOI: [10.1063/1.440249](https://doi.org/10.1063/1.440249).
- [41] J. Liebert, A. Y. Chaou, and C. Schilling. Refining and relating fundamentals of func-

- tional theory. *J. Chem. Phys.*, 158(21): 214108, 2023. DOI: [10.1063/5.0143657](https://doi.org/10.1063/5.0143657).
- [42] C. Schilling and S. Pittalis. Ensemble Reduced Density Matrix Functional Theory for Excited States and Hierarchical Generalization of Pauli’s Exclusion Principle. *Phys. Rev. Lett.*, 127:023001, 2021. DOI: [10.1103/PhysRevLett.127.023001](https://doi.org/10.1103/PhysRevLett.127.023001).
- [43] J. Liebert, F. Castillo, J.-P. Labbé, and C. Schilling. Foundation of one-particle reduced density matrix functional theory for excited states. *J. Chem. Theory Comput.*, 18(1):124–140, 2022. DOI: [10.1021/acs.jctc.1c00561](https://doi.org/10.1021/acs.jctc.1c00561).
- [44] J. Liebert. Reduced density matrix functional theory for bosons: Foundations and applications. Master’s thesis, LMU Munich, 2021. URL <https://doi.org/10.48550/arXiv.2205.02635>.
- [45] J. Liebert and C. Schilling. Deriving density-matrix functionals for excited states. *SciPost Phys.*, 14:120, 2023. DOI: [10.21468/SciPostPhys.14.5.120](https://doi.org/10.21468/SciPostPhys.14.5.120).
- [46] J. Liebert and C. Schilling. An exact one-particle theory of bosonic excitations: from a generalized Hohenberg–Kohn theorem to convexified N-representability. *New J. Phys.*, 25(1):013009, 2023. DOI: [10.1088/1367-2630/acb006](https://doi.org/10.1088/1367-2630/acb006).
- [47] F. Schautz and C. Filippi. Optimized Jastrow–Slater wave functions for ground and excited states: Application to the lowest states of ethene. *J. Chem. Phys.*, 120(23):10931–10941, 2004. DOI: [10.1063/1.1752881](https://doi.org/10.1063/1.1752881).
- [48] F. Schautz, F. Buda, and C. Filippi. Excitations in photoactive molecules from quantum monte carlo. *J. Chem. Phys.*, 121(12): 5836–5844, 2004. DOI: [10.1063/1.1777212](https://doi.org/10.1063/1.1777212).
- [49] C. Filippi, M. Zaccheddu, and F. Buda. Absorption Spectrum of the Green Fluorescent Protein Chromophore: A Difficult Case for ab Initio Methods? *J. Chem. Theory Comput.*, 5(8):2074–2087, 2009. DOI: [10.1021/ct900227j](https://doi.org/10.1021/ct900227j).
- [50] R. McWeeny. SCF theory for excited states: I. optimal orbitals for the states of a configuration. *Mol. Phys.*, 28(5):1273–1282, 1974. DOI: [10.1080/00268977400102581](https://doi.org/10.1080/00268977400102581).
- [51] E. S Sachs, J. Hinze, and N. H. Sabelli. MC-SCF calculations for six states of NaH. *J. Chem. Phys.*, 62(9):3367–3376, 1975. DOI: [10.1063/1.430989](https://doi.org/10.1063/1.430989).
- [52] M. P. Deskevich, D. J. Nesbitt, and H.-J. Werner. Dynamically weighted multiconfiguration self-consistent field: Multistate calculations for $F + H_2O \rightarrow HF + OH$ reaction paths. *J. Chem. Phys.*, 120(16):7281–7289, 2004. DOI: [10.1063/1.1667468](https://doi.org/10.1063/1.1667468).
- [53] C. L. Benavides-Riveros, L. Chen, C. Schilling, S. Mantilla, and S. Pittalis. Excitations of quantum many-body systems via purified ensembles: A unitary-coupled-cluster-based approach. *Phys. Rev. Lett.*, 129:066401, Aug 2022. DOI: [10.1103/PhysRevLett.129.066401](https://doi.org/10.1103/PhysRevLett.129.066401).
- [54] S. Matsika. Electronic structure methods for the description of nonadiabatic effects and conical intersections. *Chem. Rev.*, 121(15):9407–9449, 2021. DOI: [10.1021/acs.chemrev.1c00074](https://doi.org/10.1021/acs.chemrev.1c00074).
- [55] K. M. Nakanishi, K. Mitarai, and K. Fujii. Subspace-search variational quantum eigensolver for excited states. *Phys. Rev. Res.*, 1:033062, Oct 2019. DOI: [10.1103/PhysRevResearch.1.033062](https://doi.org/10.1103/PhysRevResearch.1.033062).
- [56] O. Higgott, D. Wang, and S. Brierley. Variational Quantum Computation of Excited States. *Quantum*, 3:156, 2019. DOI: [10.22331/q-2019-07-01-156](https://doi.org/10.22331/q-2019-07-01-156).
- [57] M. Cerezo, K. Sharma, A. Arrasmith, and P. J. Coles. Variational quantum state eigensolver. *npj Quantum Inf.*, 8(1):113, 2022. DOI: [10.1038/s41534-022-00611-6](https://doi.org/10.1038/s41534-022-00611-6).
- [58] K. M. Nakanishi, K. Mitarai, and K. Fujii. Subspace-search variational quantum eigensolver for excited states. *Phys. Rev. Res.*, 1:033062, 2019. DOI: [10.1103/PhysRevResearch.1.033062](https://doi.org/10.1103/PhysRevResearch.1.033062).
- [59] S. Yalouz, B. Senjean, J. Günther, F. Buda, T. E. O’Brien, and L. Visscher. A state-averaged orbital-optimized hybrid quantum–classical algorithm for a democratic description of ground and excited states. *Quantum Sci. Technol.*, 6(2):024004, 2021. DOI: [10.1088/2058-9565/abd334](https://doi.org/10.1088/2058-9565/abd334).
- [60] S. Yalouz, E. Koridon, B. Senjean, B. Lاسorne, G. Buda, and L. Visscher. Analytical nonadiabatic couplings and gradients within the state-averaged orbital-optimized variational quantum eigensolver. *J. Chem. The-*

- ory Comput., 18(2):776–794, 2022. DOI: [10.1021/acs.jctc.1c00995](https://doi.org/10.1021/acs.jctc.1c00995).
- [61] G. Xu, Y. B. Guo, X. Li, K. Wang, Z. Fan, Z. S. Zhou, H. J. Liao, and T. Xiang. Concurrent quantum eigensolver for multiple low-energy eigenstates. *Phys. Rev. A*, 107:052423, 2023. DOI: [10.1103/PhysRevA.107.052423](https://doi.org/10.1103/PhysRevA.107.052423).
- [62] C.-L. Hong, L. Colmenarez, L. Ding, C. L. Benavides-Riveros, and C. Schilling. Quantum parallelized variational quantum eigensolvers for excited states. *arXiv preprint arXiv:2306.11844*, 2023. DOI: [10.48550/arXiv.2306.11844](https://doi.org/10.48550/arXiv.2306.11844).
- [63] Carlos L Benavides-Riveros, Yuchen Wang, Samuel Warren, and David A Mazziotti. Quantum simulation of excited states from parallel contracted quantum eigensolvers. *arXiv:2311.05058*, 2023. DOI: [10.1088/1367-2630/ad2d1d](https://doi.org/10.1088/1367-2630/ad2d1d).
- [64] W. Kohn. Two applications of the variational method to quantum mechanics. *Phys. Rev.*, 71:635–637, 1947. DOI: [10.1103/PhysRev.71.635](https://doi.org/10.1103/PhysRev.71.635).
- [65] K. Fan. On a theorem of Weyl concerning eigenvalues of linear transformations I. *PNAS*, 35(11):652–655, 1949. DOI: [10.1073/pnas.35.11.652](https://doi.org/10.1073/pnas.35.11.652).
- [66] I. Bengtsson, Å. Ericsson, M. Kuś, W. Tadej, and K. Życzkowski. Birkhoff’s polytope andunistochastic matrices, $N = 3$ and $N = 4$. *Communications in mathematical physics*, 259:307–324, 2005. DOI: [10.1007/s00220-005-1392-8](https://doi.org/10.1007/s00220-005-1392-8).
- [67] A. W Marshall, I. Olkin, and B. C. Arnold. *Inequalities: theory of majorization and its applications*. Springer, 1979.
- [68] C. Dunkl and K. Życzkowski. Volume of the set of unistochastic matrices of order 3 and the mean Jarlskog invariant. *J. Math. Phys.*, 50(12):123521, 2009. DOI: [10.1063/1.3272543](https://doi.org/10.1063/1.3272543).
- [69] G. Birkhoff. Three observations on linear algebra. *Univ. Nac. Tacuman, Rev. Ser. A*, 5:147–151, 1946. URL <https://cir.nii.ac.jp/crid/1573387450959988992>.
- [70] J. Von Neumann. A certain zero-sum two-person game equivalent to the optimal assignment problem. *Contributions to the Theory of Games*, 2(0):5–12, 1953. DOI: [10.1515/9781400881970-002](https://doi.org/10.1515/9781400881970-002).
- [71] M. Marcus and H. Minc. *A survey of matrix theory and matrix inequalities*, volume 14. Courier Corporation, 1992.
- [72] R. A. Brualdi and P. M. Gibson. Convex polyhedra of doubly stochastic matrices. I. Applications of the permanent function. *J. Comb. Theory Ser. A.*, 22(2):194–230, 1977. DOI: [10.1016/0097-3165\(77\)90051-6](https://doi.org/10.1016/0097-3165(77)90051-6).
- [73] A. Postnikov. Permutohedra, associahedra, and beyond. *Int. Math. Res. Not.*, 2009(6):1026–1106, 2009. DOI: [doi:10.1093/imrn/rnn153](https://doi.org/10.1093/imrn/rnn153).
- [74] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge university press, 1952.
- [75] Ö. Legeza and G. Fáth. Accuracy of the density-matrix renormalization-group method. *Phys. Rev. B*, 53:14349–14358, 1996. DOI: [10.1103/PhysRevB.53.14349](https://doi.org/10.1103/PhysRevB.53.14349).
- [76] G. K.-L. Chan and M. Head-Gordon. Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group. *J. Chem. Phys.*, 116(11):4462–4476, 2002. DOI: [10.1063/1.1449459](https://doi.org/10.1063/1.1449459).
- [77] Ö. Legeza, J. Röder, and B. A. Hess. Controlling the accuracy of the density-matrix renormalization-group method: The dynamical block state selection approach. *Phys. Rev. B*, 67:125114, 2003. DOI: [10.1103/PhysRevB.67.125114](https://doi.org/10.1103/PhysRevB.67.125114).
- [78] A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and Jan von Delft. Variational matrix-product-state approach to quantum impurity models. *Phys. Rev. B*, 80:165117, 2009. DOI: [10.1103/PhysRevB.80.165117](https://doi.org/10.1103/PhysRevB.80.165117).
- [79] K. H. Marti and M. Reiher. Dmrg control using an automated richardson-type error protocol. *Mol. Phys.*, 108(3-4):501–512, 2010. DOI: [10.1080/00268971003657078](https://doi.org/10.1080/00268971003657078).
- [80] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan. The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.*, 142(3):034102, 2015. DOI: [10.1063/1.4905329](https://doi.org/10.1063/1.4905329).
- [81] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haege-man. Variational optimization algorithms for uniform matrix product states.

- Phys. Rev. B*, 97:045145, 2018. DOI: [10.1103/PhysRevB.97.045145](https://doi.org/10.1103/PhysRevB.97.045145).
- [82] C. Hubig, J. Haegeman, and U. Schollwöck. Error estimates for extrapolations with matrix-product states. *Phys. Rev. B*, 97:045125, 2018. DOI: [10.1103/PhysRevB.97.045125](https://doi.org/10.1103/PhysRevB.97.045125).
- [83] A. A. Holmes, N. M. Tubman, and C. J. Umrigar. Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J. Chem. Theory Comput.*, 12(8):3674–3680, 2016. DOI: [10.1021/acs.jctc.6b00407](https://doi.org/10.1021/acs.jctc.6b00407).
- [84] H. G. A. Burton and P.-F. Loos. Rationale for the extrapolation procedure in selected configuration interaction. *arXiv preprint arXiv:2312.12530*, 2023. DOI: [10.48550/arXiv.2312.12530](https://doi.org/10.48550/arXiv.2312.12530).
- [85] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.*, 5(1):4213, 2014. DOI: [10.1038/ncomms5213](https://doi.org/10.1038/ncomms5213).
- [86] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. Akash-Narayanan, A. Asadi, J. Miguel A., U. Azad, S. Banning, C. Blank, T. R. Bromley, B. A. Cordier, J. Ceroni, A. Delgado, O. Di Matteo, A. Dusko, T. Garg, D. Guala, A. Hayes, R. Hill, A. Ijaz, T. Isacsson, D. Ittah, S. Jahangiri, P. Jain, E. Jiang, A. Khandelwal, K. Kottmann, R. A. Lang, C. Lee, T. Loke, A. Lowe, K. McKiernan, J. J. Meyer, J. A. Montañez-Barrera, R. Moyard, Z. Niu, L. J. O’Riordan, S. Oud, A. Panigrahi, C.-Y. Park, D. Polatajko, N. Quesada, C. Roberts, N. Sá, I. Schoch, B. Shi, S. Shu, S. Sim, A. Singh, I. Strandberg, J. Soni, A. Száva, S. Thabet, R. A. Vargas-Hernández, T. Vincent, N. Vitucci, M. Weber, D. Wierichs, R. Wiersema, M. Willmann, V. Wong, S. Zhang, and N. Killoran. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2022. DOI: [10.48550/arXiv.1811.04968](https://doi.org/10.48550/arXiv.1811.04968).
- [87] Maria Schuld, Alex Bocharov, Krysta M. Svore, and Nathan Wiebe. Circuit-centric quantum classifiers. *Phys. Rev. A*, 101:032308, 2020. DOI: [10.1103/PhysRevA.101.032308](https://doi.org/10.1103/PhysRevA.101.032308).
- [88] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980*, 2017. DOI: [10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980).

A Proofs of Theorems

In this section we provide the proofs to the theorems in the main text. Most of the proofs will resort to the geometric picture of the Birkhoff polytope or permutohedron. To be more specific, we will formulate the proofs in terms of a linear constrained optimization problem on a convex polytope. The linear constraint is represented by a hyperplane $\mathbb{A}(\delta)$ (defined by the equality constraint $\Delta E_{\mathbf{w}} = \delta$) that cuts through the polytope and intersects at various edges. Each intersected edges connects two vertices on the polytope, one on the negative side of $\mathbb{A}(\delta)$ ($\Delta E_{\mathbf{w}} < \delta$), which we will call a $(-)$ -vertex (in the main text this is called a reference vertex), and the other on the positive side ($\Delta E_{\mathbf{w}} > \delta$), which we will call a $(+)$ -vertex. By identifying all $(\pm)$ -vertices one effectively identifies all edges intersected by the linear constraint. The linear constrained convex optimization problem is then reduced to the convex optimization problem within the convex intersection between the constraint and the polytope.

A.1 Proof of Theorem 3

Proof. We first reformulate the statement in terms of the permutohedron $P(\mathbf{w})$. Recall that

$$\Delta \rho_{\mathbf{w}}(\tilde{\mathbf{w}}) = 2\mathbf{w}^T(\mathbf{w} - \tilde{\mathbf{w}}), \quad (64)$$

$$\Delta E_{\mathbf{w}}(\tilde{\mathbf{w}}) = (\tilde{\mathbf{w}} - \mathbf{w})^T \mathbf{E}, \quad (65)$$

where $\tilde{\mathbf{w}} \in P(\mathbf{w})$. We wish to determine the range of $\Delta \rho_{\mathbf{w}}(\tilde{\mathbf{w}})$, for a given value of $\Delta E_{\mathbf{w}}(\tilde{\mathbf{w}}) = \delta$ where $0 < \delta < g_{\mathbf{w}, \mathbf{E}}$.

In that case, one easily verifies that the only $(-)$ -vertex is $\mathbf{w}$. The type of neighboring vertices to $\mathbf{w}$ (the $(+)$ -vertices) depends on the number of unique elements in $\mathbf{w}$. If $\mathbf{w}$ is such that $w_0 > w_1 > \dots > w_{D-1}$, then there are $D-1$ neighboring vertices, namely the $D-1$ adjacent transpositions of $\mathbf{w}$. If $\mathbf{w}$ is such that $w_0 > w_1 > \dots > w_K = \dots = w_{D-1} = 0$ for some $K < D-1$, there are also $D-1$ neighboring vertices, but they consist of K adjacent transpositions $\mathbf{S}_{k,k+1}$ where $0 \leq k < K$, and $D-K$ non-adjacent transpositions $\mathbf{S}_{K-1,k+1}$ where $k \geq K$. The index tuples (k, l) such that $\mathbf{S}_{k,l}\mathbf{w}$ is a $(+)$ -vertex is conveniently grouped by the set $\mathcal{I}_{\mathbf{w}}$.

The linear plane $\mathbb{A}(\delta)$ defined by $\Delta E_{\mathbf{w}}(\tilde{\mathbf{w}}) = \delta$ intersects all edges connecting the vertex $\mathbf{w}$, at points

$$\mathbf{v}_{k,l} = (1 - p_{k,l})\mathbf{w} + p_{k,l}\mathbf{S}_{k,l}\mathbf{w}, \quad (k, l) \in \mathcal{I}_{\mathbf{w}} \quad (66)$$

where $p_{k,l} = \delta / [(w_k - w_l)(E_l - E_k)]$. Since the intersection $\mathbb{A}(\delta) \cap P(\mathbf{w})$ is again convex with extremal points being the $\mathbf{v}_{k,l}$'s with $(k, l) \in \mathcal{I}_{\mathbf{w}}$, we can deduce that for any $\mathbf{v} \in \mathbb{A} \cap P(\mathbf{w})$

$$\min_{(k,l) \in \mathcal{I}_{\mathbf{w}}} \Delta \rho_{\mathbf{w}}(\mathbf{v}_{k,l}) \leq \Delta \rho_{\mathbf{w}}(\mathbf{v}) \leq \max_{(k,l) \in \mathcal{I}_{\mathbf{w}}} \Delta \rho_{\mathbf{w}}(\mathbf{v}_{k,l}). \quad (67)$$

By substituting in the exact expression of the error of the ensemble state at the points $\mathbf{v}_{k,l}$'s

$$\Delta \rho_{\mathbf{w}}(\mathbf{v}_{k,l}) = 2\delta \frac{w_k - w_l}{E_l - E_k}, \quad (68)$$

we arrive at both the lower and upper bounds. $\square$

A.2 Proof of Theorem 4

Proof. Recall that

$$\Delta \Psi_k(\mathbf{X}) = 1 - X_{kk}, \quad (69)$$

$$\Delta E_{\mathbf{w}}(\mathbf{X}) = \mathbf{w}^T(\mathbf{X}^T - \mathbb{1})\mathbf{E}, \quad (70)$$

where $X_{kl} = |\langle \Psi_k | \tilde{\Psi}_l \rangle|^2$ is an unistochastic matrix in the Birkhoff polytope B_D . We are interested in the region where $0 < \Delta E_{\mathbf{w}} = \delta < g_{\mathbf{w}, \mathbf{E}}$.

It is not difficult to see that the lower bound for $\Delta\Psi_k$ is zero, since a finite error in the ensemble energy can be due to errors in other eigenenergies than E_k . For the upper bound, let us suppose $\mathbf{w}$ contains K positive entries. Then the $(-)$ -vertices on the Birkhoff polytope are $\mathbb{1}$ and permutations $\mathbf{S}$ that act only non-trivially on the zeros elements of $\mathbf{w}$. According to the property of the Birkhoff polytope[72], $(+)$ -vertices are permutations that differ by a single cycle from the $(-)$ -vertices. Notice that for any permutation $\mathbf{S}$, the resulting error $\Delta\Psi_k$ in the k -th eigenstate can only be 0 or 1. Therefore, to find the upper bound of $\Delta\Psi_k$, it suffices to find among the $(+)$ -vertices involving the k -th eigenstate, that realizes the smallest error in the ensemble energy (and thus maximizing the ratio $\Delta\Psi_k/\Delta E_{\mathbf{w}}$). Such smallest error is given by $\min\{g_{k-1,k}(\mathbf{w}, \mathbf{E}), g_{k,k+1}(\mathbf{w}, \mathbf{E})\}$, for $k \geq 1$, where $g_{kl}(\mathbf{w}, \mathbf{E}) \equiv (w_k - w_l)(E_l - E_k)$. When $k = 0$, the smallest error is simply $g_{0,1}(\mathbf{w}, \mathbf{E})$. We therefore conclude that

$$0 \leq \frac{\Delta\Psi_k}{\Delta E_{\mathbf{w}}} \leq \begin{cases} g_{0,1}^{-1}(\mathbf{w}, \mathbf{E}), & k = 0 \\ \max\{g_{k-1,k}^{-1}(\mathbf{w}, \mathbf{E}), g_{k,k+1}^{-1}(\mathbf{w}, \mathbf{E})\}, & k \geq 1. \end{cases} \quad (71)$$

□

A.3 Proof of Theorem 5

Proof. The proof follows similarly to the proof of Theorem 4. We assume there are exactly K positive elements in $\mathbf{w}$. The $(-)$ -vertices are still the identity matrix $\mathbb{1}$ and permutations $\mathbf{S}_-$ that acts only non-trivially on the zero elements of $\mathbf{w}$. And the $(+)$ -vertices are permutations that differ by a single cycle from the $(-)$ -vertices.

A key difference is that for a $(+)$ -vertex $\mathbf{S}_+$, the corresponding error of the K lowest eigenstates is

$$1 \leq \sum_{k=0}^{K-1} \Delta\Psi_k(\mathbf{S}_+) \leq K. \quad (72)$$

Notice that the error cannot be zero, as that can only be realized by one of the $(-)$ -vertices. When $\mathbf{w}$ is strictly decreasing ($K = D$), the lower bound is improved to 2, since the minimal number of misplaced eigenstates is 2 in that case. The exact (integer) value of the error depends on the number of non-zero elements of $\mathbf{w}$ that $\mathbf{S}_+$ non-trivially acts on. We define the following two helper quantities

$$\begin{aligned} \delta_{\mathbf{w}, \mathbf{E}}^{(1)} &= \min_{i,j < K} (w_i^\downarrow - w_j^\downarrow)(E_j^\uparrow - E_i^\uparrow), \\ \delta_{\mathbf{w}, \mathbf{E}}^{(2)} &= \min_{i < K, j \geq K} (w_i^\downarrow - w_j^\downarrow)(E_j^\uparrow - E_i^\uparrow). \end{aligned} \quad (73)$$

Lemma 1. Let $\mathbf{S}_L = \mathbf{S}_{\mathbf{n}}\mathbf{S}_-$ be a $(+)$ -vertex of the Birkhoff polytope with respect to the hyperplane $\mathbb{A}$ defined by $0 < \Delta E_{\mathbf{w}} = \delta < g_{\mathbf{w}, \mathbf{E}}$. Here, $\mathbf{S}_-$ is a $(-)$ -vertex and $\mathbf{S}_{\mathbf{n}}$ is a single L -cycle ($L \geq 2$) denoted in cycle notation as $\mathbf{n} = (n_1, n_2, \dots, n_L)$ where exactly L many w_{n_k} 's are positive and distinct ($1 \leq L' \leq L$), while the others are zero. Then we have

$$\Delta E_{\mathbf{w}}(\mathbf{S}_L) \leq \min\left(L', \left\lfloor \frac{L}{2} \right\rfloor\right) G_{\mathbf{w}, \mathbf{E}}. \quad (74)$$

When $\delta_{\mathbf{w}, \mathbf{E}}^{(1)} \leq 2\delta_{\mathbf{w}, \mathbf{E}}^{(2)}$ and $L' > 1$, we have

$$(L' - 1)\delta_{\mathbf{w}, \mathbf{E}}^{(1)} \leq \Delta E_{\mathbf{w}}(\mathbf{S}_L), \quad (75)$$

and when $\delta_{\mathbf{w}, \mathbf{E}}^{(1)} > 2\delta_{\mathbf{w}, \mathbf{E}}^{(2)}$ or $L' = 1$, we have

$$(2L' - 1)\delta_{\mathbf{w}, \mathbf{E}}^{(2)} \leq \Delta E_{\mathbf{w}}(\mathbf{S}_L). \quad (76)$$

Specially, when $L' = L$, the inequalities become

$$(L - 1)g_{\mathbf{w}, \mathbf{E}} \leq \Delta E_{\mathbf{w}}(\mathbf{S}_L) \leq \left\lfloor \frac{L}{2} \right\rfloor G_{\mathbf{w}, \mathbf{E}}. \quad (77)$$

Proof. For the upper bound, we use the rearrangement inequality

$$\begin{aligned}\Delta E_{\mathbf{w}}(\mathbf{S}_L) &\leq \sum_{l=1}^{\lfloor L/2 \rfloor} (w_{n_l^\uparrow} - w_{n_{L-l+1}^\uparrow})(E_{n_{L-l+1}^\uparrow} - E_{n_l^\uparrow}) \\ &\leq \min\left(L', \left\lfloor \frac{L}{2} \right\rfloor\right) G_{\mathbf{w}, \mathbf{E}}.\end{aligned}\tag{78}$$

The lower bound can be proven by induction. When $L = 2$, the lower bounds for $\Delta E_{\mathbf{w}}(\mathbf{S}_L)$ are trivially satisfied. Let $\mathbf{S}_{L+1} = \mathbf{S}_{\mathbf{n}}\mathbf{S}_{-}$ where $\mathbf{S}_{\mathbf{n}}$ is a $(L+1)$ -cycle with $\mathbf{n} = (n_1, n_2, \dots, n_K, n_{L+1})$ and $L \geq 2$. Using the cyclic properties of the cycle notation, we can assume that n_{L+1} is the index corresponding to the largest energy (or equivalently the smallest weight). By the assumption of the induction, the L -cycle $\mathbf{S}_L$ representing $(n_1, n_2, \dots, n_L)$ satisfies the lower bound. Then

$$\begin{aligned}&\mathbf{w}^T(\mathbf{S}_{L+1} - \mathbb{1})\mathbf{E} \\ &= \sum_{l=1}^{L-1} w_{n_l}(E_{n_{l+1}} - E_{n_l}) + w_{n_L}(E_{n_{L+1}} - E_{n_L}) \\ &\quad + w_{n_{L+1}}(E_{n_1} - E_{n_{L+1}}) \\ &= \mathbf{w}^T(\mathbf{S}_L - \mathbb{1})\mathbf{E} + (w_{n_L} - w_{n_{L+1}})(E_{n_{L+1}} - E_{n_1}).\end{aligned}\tag{79}$$

Each step of induction acquires an additional term. Notice that the additional term vanishes if and only if both w_{n_L} and $w_{n_{L+1}}$ are zero. When at least one of them is positive, we have the following inequality

$$\min(\delta_{\mathbf{w}, \mathbf{E}}^{(1)}, \delta_{\mathbf{w}, \mathbf{E}}^{(2)}) \leq (w_{n_L} - w_{n_{L+1}})(E_{n_{L+1}} - E_{n_1}).\tag{80}$$

Depending on the order of the additional index n_{L+1} at each step of the induction, which we call a line-up, the above bound can be made tighter. Namely, if two consecutive elements n_L and n_{L+1} are positive, we can conclude the lower bound of the additional term is $\delta_{\mathbf{w}, \mathbf{E}}^{(1)}$. If only one of them is positive, the lower bound becomes $\delta_{\mathbf{w}, \mathbf{E}}^{(1)}$. If we use 1 to represent a positive w_{n_l} and 0 for a vanishing one, then the line-up that minimizes the accumulated sum in the induction is one of the following two cases. When $\delta_{\mathbf{w}, \mathbf{E}}^{(1)} \geq 2\delta_{\mathbf{w}, \mathbf{E}}^{(2)}$ and $L' > 1$, the minimizing line-up is

$$\underbrace{1, 1, \dots, 1}_{L'}, 0, 0, \dots, 0.$$

The accumulated sum in the induction is then

$$(L' - 1)\delta_{\mathbf{w}, \mathbf{E}}^{(1)} \leq \Delta E_{\mathbf{w}}(\mathbf{S}_L).\tag{81}$$

When $\delta_{\mathbf{w}, \mathbf{E}}^{(1)} > 2\delta_{\mathbf{w}, \mathbf{E}}^{(2)}$ or $L' = 1$, the minimizing line-up is

$$\begin{aligned}&\underbrace{1, 0, 1, 0, \dots, 1, 0, 0, \dots, 0}_{2L'}, \quad L' \leq L/2, \\ &\underbrace{0, 1, 0, 1, \dots, 0, 1, 1, \dots, 1}_{2(L-L')}, \quad L' > L/2.\end{aligned}$$

Straightforward algebra shows that the accumulated sum in the induction is then

$$(2L' - 1)\delta_{\mathbf{w}, \mathbf{E}}^{(2)} \leq \Delta E_{\mathbf{w}}(\mathbf{S}_L).\tag{82}$$

□

From Lemma 1, we can derive that for a (+)-vertex $\mathbf{S}_L$ acting non-trivially on L' non-zeros elements of $\mathbf{w}$, the following lower bound on the error of the lowest K ($K < D-1$) eigenstates

$$\frac{\sum_{k=0}^{K-1} \Delta \Psi_k(\mathbf{S}_L)}{\Delta E_{\mathbf{w}}(\mathbf{S}_L)} \geq \frac{L'}{\min\left(L', \left\lfloor \frac{L'}{2} \right\rfloor\right) G_{\mathbf{w}, \mathbf{E}}} \geq \frac{1}{G_{\mathbf{w}, \mathbf{E}}}. \quad (83)$$

When $K = D$, we always have $L' = L$, which gives us a tighter lower bound

$$\frac{\sum_{k=0}^{D-1} \Delta \Psi_k(\mathbf{S}_L)}{\Delta E_{\mathbf{w}}(\mathbf{S}_L)} \geq \frac{2}{G_{\mathbf{w}, \mathbf{E}}}. \quad (84)$$

For the upper bound, we again separate the two cases. According to Lemma 1, when $\delta_{\mathbf{w}, \mathbf{E}}^{(1)} \leq 2\delta_{\mathbf{w}, \mathbf{E}}^{(2)}$ and $L' > 1$

$$\frac{\sum_{k=0}^{K-1} \Delta \Psi_k(\mathbf{S}_L)}{\Delta E_{\mathbf{w}}(\mathbf{S}_L)} \leq \frac{L'}{(L' - 1)\delta_{\mathbf{w}, \mathbf{E}}^{(1)}} \leq \frac{2}{\delta_{\mathbf{w}, \mathbf{E}}^{(1)}}, \quad (85)$$

where equality is achieved when $L' = 2$. When $\delta_{\mathbf{w}, \mathbf{E}}^{(1)} > 2\delta_{\mathbf{w}, \mathbf{E}}^{(2)}$ or $L' = 1$,

$$\frac{\sum_{k=0}^{K-1} \Delta \Psi_k(\mathbf{S}_L)}{\Delta E_{\mathbf{w}}(\mathbf{S}_L)} \leq \frac{L'}{(2L' - 1)\delta_{\mathbf{w}, \mathbf{E}}^{(2)}} \leq \frac{1}{\delta_{\mathbf{w}, \mathbf{E}}^{(2)}}, \quad (86)$$

where equality is achieved when $L' = 1$. Specially, when $K = D$, we always have $L' = L \geq 2$. Therefore

$$\frac{\sum_{k=0}^{D-1} \Delta \Psi_k(\mathbf{S}_L)}{\Delta E_{\mathbf{w}}(\mathbf{S}_L)} \frac{L}{(L - 1)g_{\mathbf{w}, \mathbf{E}}} \leq \frac{2}{g_{\mathbf{w}, \mathbf{E}}}. \quad (87)$$

□

A.4 Proof of Theorem 7

Proof. Let us first rewrite the ensemble energy error as

$$\begin{aligned} \Delta E_{\mathbf{w}} &= \sum_{k=0}^{D-1} (w_k - w_{k+1}) \sum_{j=0}^k (\tilde{E}_j - E_j) \\ &= \sum_{k=0}^{D-2} (w_k - w_{k+1}) F_k, \end{aligned} \quad (88)$$

where $w_D \equiv 0$ and $F_k = \sum_{j=0}^k (\tilde{E}_j - E_j)$. F_k are non-negative thanks to Theorem 1, and specially $F_{D-1} = \text{Tr}[\hat{H} - \hat{H}] = 0$. If $\mathbf{w}$ is strictly decreasing, then (defining $F_{-1} = 0$)

$$\begin{aligned} \sum_{k=0}^{D-1} |\Delta E_k| &= \sum_{k=0}^{D-1} |F_k - F_{k-1}| \leq 2 \sum_{k=0}^{D-1} F_k \\ &\leq \frac{2}{\min_{k < D-1} (w_k - w_{k+1})} \sum_{k=0}^{D-2} (w_k - w_{k+1}) F_k \\ &= \frac{2\Delta E_{\mathbf{w}}}{\min_{k < D-1} (w_k - w_{k+1})}. \end{aligned} \quad (89)$$

For the lower bound, we first define the positive and negative parts (both as positive values) $\Delta\tilde{\mathbf{E}}_{\pm}$ of the vector $\Delta\tilde{\mathbf{E}} = (\Delta E_k)_{k=0}^{D-1}$, where the negative and positive entries are set to zero, respectively.

$$\begin{aligned}\Delta E_{\mathbf{w}} &= \sum_{k=0}^{D-1} w_k [(\Delta E_+)_{k-1} - (\Delta E_-)_k] \\ &\leq w_0 \sum_{k=0}^{D-1} (\Delta E_+)_{k-1} - w_{D-1} \sum_{k=0}^{D-1} (\Delta E_-)_k \\ &= \frac{w_0 - w_{D-1}}{2} \sum_{k=0}^{D-1} |\Delta E_k|. \end{aligned} \quad (90)$$

In the last line we used the identity (recall that $\sum_{k=0}^{D-1} (\Delta E_+)_{k-1} - \sum_{k=0}^{D-1} (\Delta E_-)_k = F_{D-1} = 0$)

$$\sum_{k=0}^{D-1} (\Delta E_+)_{k-1} = \frac{1}{2} \sum_{k=0}^{D-1} |\Delta E_k|. \quad (91)$$

If $w_{K-1} > w_K > w_{K+1}$ for some $0 < K < D-1$, then the total error of the lowest K eigenenergies are bounded from above similarly as

$$\begin{aligned} \sum_{k=0}^{K-1} |\Delta E_k| &= \sum_{k=0}^{K-1} |F_k - F_{k-1}| \leq 2 \sum_{k=0}^{K-2} F_k + F_{K-1} \\ &\leq \Delta E_{\mathbf{w}} \max_{k < K} \left\{ \frac{2\Theta(K-k-1)}{w_k - w_{k+1}} \right\}. \end{aligned} \quad (92)$$

For the lower bound, we simply observe that

$$\tilde{E}_{\mathbf{w}} \leq \sum_{k=0}^{K-1} w_k |\Delta E_k| \leq w_0 \sum_{k=0}^{K-1} |\Delta E_k|, \quad (93)$$

which completes the proof. $\square$

B Derivation for optimal weight vectors

Recall the μ -notation for parameterizing a non-increasing weight vector $\mathbf{w}$

$$\begin{aligned} \mu_k &= w_k - w_{k+1}, \quad k < D-1, \\ \mu_{D-1} &= w_{D-1} \end{aligned} \quad (94)$$

where μ_k 's satisfy

$$\begin{aligned} 0 &\leq \mu_k \leq \frac{1}{k+1}, \\ \sum_{k=0}^{D-1} (k+1) \mu_k &= 1. \end{aligned} \quad (95)$$

Now we would like to solve the following minimization for the linear upper bound of the error of all eigenenergies $\sum_{k=0}^{D-1} |\Delta E_k|$

$$\min_{\boldsymbol{\mu}} \max_{k < D-1} \frac{1}{\mu_k}. \quad (96)$$

Notice that μ_{D-1} does not enter the above expression. Therefore, we can reformulate the minimization with only inequality constraints on $(\mu_k)_{k=0}^{D-2}$. To be explicit, the minimization becomes

$$\min_{0 \leq \mu_k < D-1 \leq 1/(k+1)} \max_{k < D-1} \mu_k^{-1}. \quad (97)$$

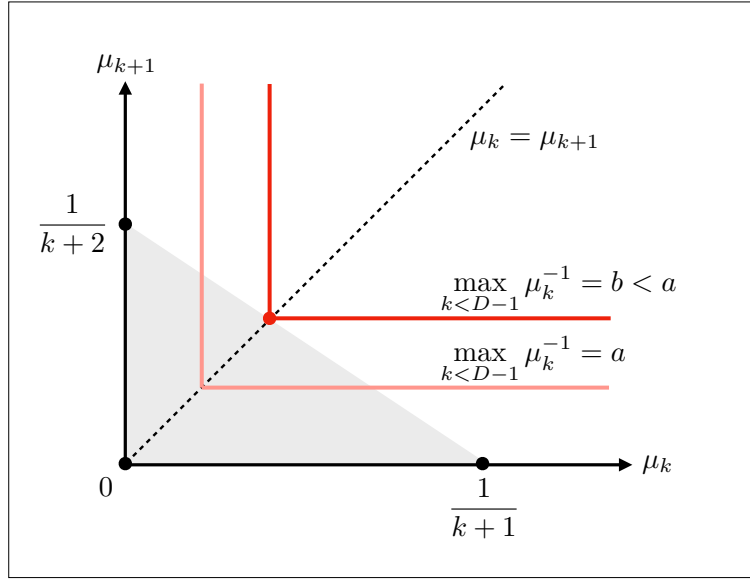


Figure 9: Pictorial illustration of the derivation for optimal weights. See text for more details.

In Figure 9 we illustrate how the minimum is obtained. The gray shaded area is the domain of minimization of $\boldsymbol{\mu}$. The equation $\max_{k < D-1} \mu_k^{-1} = a$ is an L-shaped curve (in color red) with its corner on the line defined by $\mu_0 = \mu_1 = \dots = \mu_{D-2}$. As the curve moves away from the origin along the line $\mu_0 = \mu_1 = \dots = \mu_{D-2}$, the value of $\max_{k < D-1} \mu_k^{-1}$ decreases. It is then clear that the minimum is obtained as the curve touches the slanted boundary of the gray domain, defined by $\mu_0 + 2\mu_1 + \dots + (D-1)\mu_{D-2} = 1$ and $\mu_{D-1} = 0$. Solving its intersection with the line $\mu_0 = \mu_1 = \dots = \mu_{D-2}$, we obtain the optimal weight vector

$$\boldsymbol{w} = \frac{2}{D(D-1)}(D-1, D-2, \dots, 1). \quad (98)$$

When we consider the upper bound for the error of the lowest K eigenstates, even more elements of $\boldsymbol{\mu}$ are free and do not enter the minimization

$$\min_{0 \leq \mu_k < K \leq 1/(k+1)} \max_{k < K} \frac{\Theta(K-k-1)}{\mu_k}. \quad (99)$$

Following the same logic, in the optimal solution the elements $\mu_K, \mu_{K+1}, \dots, \mu_{D-1}$ will be 0. This time, the corner of the curve defined by $\max_{k < K} \mu_k^{-1} \Theta(K-k-1) = a$ lies on the line $\mu_0 = \mu_1 = \dots = 2\mu_{K-1}$. Combining these condition would then lead to the optimal weights

$$\boldsymbol{w} = \frac{1}{K^2}(2K-1, 2K-3, \dots, 1). \quad (100)$$

As for the optimal weights for the error of the eigenstates, the proof follows similarly. The only difference is that weight gaps μ_k 's are modified by the according energy gaps. For example, for targeting all eigenstates, the optimal weight vector is the minimizer of the following problem

$$\min_{0 \leq \mu_k < D-1 \leq 1/(k+1)} \max_{k < D-1} \mu_k^{-1} (E_{k+1} - E_k)^{-1}. \quad (101)$$

The solution is then given by the intersection of the following conditions

$$\begin{aligned} \mu_0 + 2\mu_1 + \dots + (D-1)\mu_{D-2} &= 1, \\ \mu_{D-1} &= 0, \\ \frac{\mu_0}{E_1 - E_0} &= \frac{\mu_1}{E_2 - E_1} = \dots = \frac{\mu_{D-2}}{E_{D-1} - E_{D-2}}. \end{aligned} \quad (102)$$

The first two conditions determine the boundary of the domain, whereas the last condition determines the point on the boundary at which the corner of the curve defined by $\max_{k < D-1} \mu_k^{-1}(E_{k+1} - E_k)^{-1} = a$ touches. For targeting the lowest K eigenstates, we simply sent all $w_{k \geq K}$ to 0, and replace $E_{D-1} - E_{D-2}$ with $(E_{D-1} - E_{D-2})/2$.