Energy error bars in direct configuration interaction iteration sequence

Zsuzsanna Tóth and Ágnes Szabados

Articles you may be interested in

Configuration interaction wave functions: A seniority number approach
J. Chem. Phys. **140**, 234103 (2014); 10.1063/1.4882881

Toward large scale vibrational configuration interaction calculations
J. Chem. Phys. **131**, 124129 (2009); 10.1063/1.3243862

Structure of the exact wave function. V. Iterative configuration interaction method for molecular systems within finite basis
J. Chem. Phys. **117**, 9 (2002); 10.1063/1.1487830

Structure of the exact wave function. IV. Excited states from exponential ansatz and comparative calculations by the iterative configuration interaction and extended coupled cluster theories
J. Chem. Phys. **116**, 1811 (2002); 10.1063/1.1430741

Subspace iteration techniques for the calculation of resonances using complex symmetric Hamiltonians
J. Chem. Phys. **112**, 2106 (2000); 10.1063/1.480778
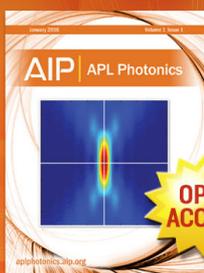
CrossMark
click for updates

Energy error bars in direct configuration interaction iteration sequence

Zsuzsanna Tóth and Ágnes Szabados[a)]
Laboratory of Theoretical Chemistry, Institute of Chemistry, Loránd Eötvös University, 1518 Budapest,
P.O. Box 32, Hungary

A computational scheme for approximate lower bound to eigenvalues of linear operators is elaborated, based on Löwdin's bracketing function. Implementation in direct full configuration interaction algorithm is presented, generating essentially just input–output increase. While strict lower bound property is lost due to approximations, test calculations result lower bounds of the same order of magnitude, as the usual upper bound, provided by the expectation value. Difference of upper and lower bounds gives an error bar, characterizing the wavefunction at the given iteration step. © 2015 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4928977]

I. INTRODUCTION

While upper bounds to the energy levels of molecular systems are harnessed in variational quantum chemical methods, lower bounds are comparatively marginal. Lower bound methods, based on the second moment, or variance of the Hamiltonian[1–5] are perhaps the most known. Some other approaches are based on operators bounding the Hamiltonian from below,[6–9] Padé approximants,[10] local energy,[11] spectral consideration,[12] or matrix representation of the Hamiltonian.[13] Eigenvalues of the two-electron reduced Hamiltonian also provide strict approximation to the exact eigenvalues from below,[14] and this property is usually conserved when N-representability conditions are partially incorporated within a reduced density matrix approach.[15]

Studies on the bracketing function

$$f(\varepsilon) = \varepsilon + \langle\varphi|(H - \varepsilon)^{-1}|\varphi\rangle^{-1} \qquad (1)$$

were initiated by Löwdin[16] and were pursued for some time[8,17–19,39] to become essentially abandoned to our date, with rare exceptions.[20] The bracketing function gives strict lower bound to an eigenvalue of the Hamiltonian, H, provided that the reference function, φ, is a normalized approximation to the state in question and ε is an appropriate upper bound. It has been recently shown,[21] that a variational principle applies for $f(\varepsilon)$, i.e., the bracketing function is stationary with respect to small variations of the reference, if and only if φ is an exact eigenfunction of H. In view of this, the inverse method of Scrinzi[22] can be regarded a variation technique for the bracketing function, with a specific choice made for the reference function, φ. We note for completeness that the inverse method is equivalent to the generalized variance method of Marmorino and closely related to Temple's lower bound.[23]

In the present work, we wish to show that Löwdin's bracketing function can be utilized at negligible extra cost in a direct configuration interaction (CI) algorithm to compute reasonably tight, approximate lower bound to the matrix

eigenvalue. Supplemented with the expectation value as upper bound, an energy error bar is obtained, corresponding to the reference function. Error bars for energy differences can be derived by subtracting appropriate bounds for the individual levels.

II. THEORY

Let us focus on the ground state and assume that the overlap of the normalized reference function φ and the exact eigenfunction ψ is nonzero (nonsingular approximation),

$$\langle\varphi|\psi\rangle \neq 0.$$

Evaluation of Eq. (1) via explicitly expressing the resolvent or Green-operator

$$G = (H - \varepsilon)^{-1} \qquad (2)$$

is computationally prohibitive. It is viable on the other hand to evaluate the effect of H on a reference function. Efficient calculation of $H|\varphi\rangle$ is the key idea behind the direct CI toolkit, initiated by Roos and Siegbahn.[24] This encourages to use a power series approximation, following a perturbative partitioning of the Hamiltonian as

$$H = H^0 + V.$$

Introducing

$$G^0 = (H^0 - \varepsilon)^{-1}$$

for the resolvent associated with the zero-order Hamiltonian, a Taylor-series expansion (also referred to as the Neumann-series[25]) of the inverse in Eq. (2) takes the form

$$G = G^0 \sum_{\nu=0}^{\infty} (-VG^0)^{\nu}. \qquad (3)$$

Convergence of the series depends on the norm of operator VG^0, in particular $\|VG^0\| < 1$ is a sufficient criterion.[25,26] From this point of view, the closer H^0 lies to H the better. At the same time the zero-order Hamiltonian must be simple enough to ensure that calculation of $G^0|\varphi\rangle$ is computationally manageable. A zero-order Hamiltonian, defined via its matrix

---

[a)]Electronic address: szabados@chem.elte.hu, Tel.: +36-1-3722500.

elements as

$$\langle\varphi|H^0|\varphi\rangle = \langle\varphi|H|\varphi\rangle,$$
$$\langle\varphi|H^0|\varphi_i\rangle = \langle\varphi|H|\varphi_i\rangle, \quad i = 1,\ldots,N,$$
$$\langle\widetilde{\varphi}_i|H^0|\varphi\rangle = \langle\widetilde{\varphi}_i|H|\varphi\rangle, \quad i = 1,\ldots,N, \quad (4)$$
$$\langle\widetilde{\varphi}_i|H^0|\varphi_i\rangle = \langle\widetilde{\varphi}_i|H|\varphi_i\rangle, \quad i = 1,\ldots,N,$$
$$\langle\widetilde{\varphi}_i|H^0|\varphi_j\rangle = 0, \quad i,j = 1,\ldots,N \wedge i \neq j,$$

achieves a fortunate balance between the two contradictory criteria. We now proceed to define φ_i and $\widetilde{\varphi}_i$ appearing above (see Eqs. (6) and (7) for their explicit expression).

Excited functions, φ_i, together with the reference, φ, span the $(N + 1)$ dimensional full CI space. Assuming a determinantal expansion of the reference as

$$|\varphi\rangle = |\xi_0\rangle x_0 + \sum_{i=1}^{N}|\xi_i\rangle x_i, \quad (5)$$

excited functions are obtained by Gram-Schmidt orthogonalization

$$|\varphi_i\rangle = |\xi_i\rangle - |\varphi\rangle x_i, \quad i = 1,\ldots,N. \quad (6)$$

Unless the associated x_0 is zero, choice for the pivot determinant, ξ_0, is arbitrary. Functions, $\widetilde{\varphi}_i$, expressed as

$$\langle\widetilde{\varphi}_i| = \langle\xi_i| - \frac{x_i}{x_0}\langle\xi_0|, \quad i = 1,\ldots,N \quad (7)$$

constitute the reciprocal set to $\{\varphi_i\}_{i=1}^N$. It is easy to check that

$$\langle\widetilde{\varphi}_i|\varphi_j\rangle = \delta_{ij}, \quad i,j = 1,\ldots,N \quad (8)$$

indeed holds.[27,28] As an alternative to the biorthogonal formulation, Löwdin orthogonalization of the set $\{\varphi_i\}_{i=1}^N$ could be performed. Expression of the Löwdin orthogonalized functions is also available in closed form.[29,30] The reason for preferring the biorthogonal treatment is the compactness of the determinantal expansion of $\widetilde{\varphi}_i$.

The zero-order Hamiltonian of Eq. (4), illustrated in Fig. 1, neglects any interaction among excited states but includes all interaction between excited states and the reference via the Hamiltonian. This choice for H^0 is essentially the approximation designated A_0 by Shavitt,[31] formulated here in a biorthogonal manner. The motivation behind this partitioning is to include all matrix elements which contribute to a first order correction to φ in an Epstein-Nesbet (EN) partitioning.[32,33] In fact, the eigenvector of Eq. (4) matches with the first-order Brillouin-Wigner (BW) wavefunction in EN partitioning and the corresponding eigenvalue is of order 2 in BW theory.[31] Since Eq. (4) includes all direct interaction with the reference, $H^0|\varphi\rangle = H|\varphi\rangle$ is fulfilled once φ becomes an exact eigenvector. As a consequence $G^0|\varphi\rangle = G|\varphi\rangle$ holds, hence only the first term of Eq. (3) contributes. Let us note here that definition Eq. (4) shares a downside with EN partitioning: it is not invariant to unitary transformation among φ_i.[34]

An important characteristic of the present approach is that series Eq. (3), even if summed to infinite order, produces the inverse of matrix $H - \varepsilon$, instead of the matrix representation of the genuine operator inverse. As a consequence, function $f(\varepsilon)$, with appropriate value for ε, approximates *matrix eigenvalues* from below. Such an approximation may still be an upper bound to the true eigenvalue, the so-called complete CI limit.
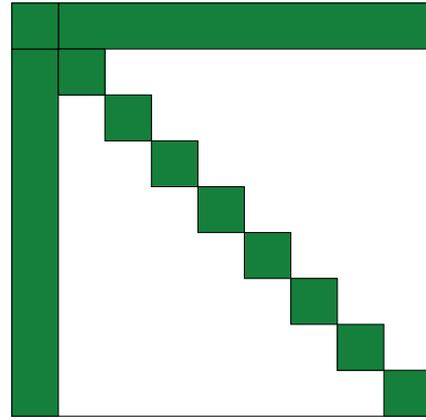

FIG. 1. Schematic representation of the zero-order Hamiltonian of Eq. (4), cf. Shavitt's A_0 approximation.[31]

Due to associated conceptual and numerical difficulties, lower bounds to true eigenvalues of the Hamiltonian are typically evaluated for simple quantum systems, e.g., the He atom.[5,6,9,11] In the present study we rest satisfied with the aim of bounding full CI eigenvalues but not those of the complete CI limit. In return, an algorithm computable for any quantum chemical system is devised.

A. The zero-order resolvent

The simple structure of matrix $\boldsymbol{H^0}$ defined in Eq. (4) facilitates to express $\boldsymbol{G^0}$ explicitly. Introducing shorthands $\varphi_0 \equiv \varphi$ and $H_{\widetilde{i}j} = \langle\widetilde{\varphi}_i|H|\varphi_j\rangle$, the determinant of $(\boldsymbol{H^0} - \boldsymbol{\varepsilon})$ reads

$$\det(\boldsymbol{H^0} - \boldsymbol{\varepsilon}) = \eta \prod_{i=1}^{N}(H_{\widetilde{i}i} - \varepsilon), \quad (9)$$

with

$$\eta = H_{00} - \sum_{i=1}^{N}\frac{H_{0i}H_{\widetilde{i}0}}{H_{\widetilde{i}i} - \varepsilon} - \varepsilon. \quad (10)$$

The top left matrix element of $\boldsymbol{G^0}$ is given by $G^0_{00} = \eta^{-1}$; elements in the zeroth column take the form

$$G^0_{\widetilde{i}0} = -\eta^{-1}\frac{H_{\widetilde{i}0}}{H_{\widetilde{i}i} - \varepsilon}, \quad i = 1,\ldots,N \quad (11)$$

and the zeroth row similarly reads as

$$G^0_{0i} = -\eta^{-1}\frac{H_{0i}}{H_{\widetilde{i}i} - \varepsilon}, \quad i = 1,\ldots,N. \quad (12)$$

Matrix elements for $i,j \neq 0$ look

$$G^0_{\widetilde{i}j} = \delta_{ij}(H_{\widetilde{i}i} - \varepsilon)^{-1}$$
$$+ \eta^{-1}\frac{H_{\widetilde{i}0}H_{0j}}{(H_{\widetilde{i}i} - \varepsilon)(H_{\widetilde{j}j} - \varepsilon)}, \quad i,j = 1,\ldots,N. \quad (13)$$

Note that $G^0_{\widetilde{i}j}$ in Eq. (13) is factorized in indices i and j. Consequently, the cost associated with the evaluation of the effect of $\boldsymbol{G^0}$ on a trial vector is a couple of scalar products. Regarding central processing unit (CPU) time, this is negligible as compared to the transformation by H.

Starting from a column vector $\boldsymbol{q}$, associated with a function

$$\sum_{i=0}^{N} |\xi_i\rangle q_i, \tag{14}$$

we give here the effect of $\boldsymbol{G^0}$ for completeness

$$\boldsymbol{G^0 q} = \boldsymbol{x}\,(\eta^{-1}A - B) + \boldsymbol{b}, \tag{15}$$

with $\boldsymbol{x}$ denoting the column vector formed of x_i of Eq. (5). Scalar A is given by

$$A = \sum_{i=0}^{N} x_i q_i - \sum_{i=1}^{N} \frac{H_{\widetilde{0}i}}{H_{\widetilde{i}i} - \varepsilon} \left(q_i - x_i \frac{q_0}{x_0} \right). \tag{16}$$

Components of vector $\boldsymbol{b}$ read

$$b_i = \frac{1}{H_{\widetilde{i}i} - \varepsilon} \left\{ q_i - x_i \frac{q_0}{x_0} - \eta^{-1} A H_{\widetilde{i}0} \right\}, \quad i = 1,\dots,N, \tag{17}$$

with $b_0 = 0$. Finally,

$$B = \sum_{i=1}^{N} x_j b_j. \tag{18}$$

B. Projecting into the Davidson-subspace

Evaluation of the terms of Eq. (3) to any order beyond the zeroth implies acting with H apart from G^0. This is the rate determining step of the algorithm, which has been performed already n times at iteration step n. For this reason projector $\boldsymbol{O_D}$ is introduced as

$$\boldsymbol{O_D} = \sum_{i=1}^{n} \boldsymbol{c^i c^{iT}}, \tag{19}$$

with $\boldsymbol{c^i}$ being the ith among basis (column) vectors spanning the Davidson-subspace,[35] and $\boldsymbol{c^{iT}}$ its transpose. To reduce computational cost, we experiment with inserting O_D preceding occurrences of $V = H - H^0$ in Eq. (3). The set $\{\boldsymbol{Hc^i}\}_{i=1}^{n}$ being available on disk, the effect of HO_D on a trial function of Eq. (14) is comparatively easily obtained as

$$\boldsymbol{HO_D q} = \sum_{i=1}^{n} \boldsymbol{Hc^i} \sum_{j=0}^{N} c_j^i q_j. \tag{20}$$

C. Analysis in the vicinity of the exact wavefunction

Let us now investigate the dependence of the bracketing function on a small error in the trial function. For this end let us suppose that ψ_i are exact, normalized eigenvectors of H and write

$$|\varphi\rangle = |\psi_0\rangle \mu_0 + |\chi\rangle \mu, \tag{21}$$

with $\langle \psi_0 | \chi \rangle = 0$. Assuming for simplicity, that the error in φ, function χ is normalized, we have $\mu_0^2 + \mu^2 = 1$. Without invoking any approximation, the bracketing function of Eq. (1) depends on μ as

$$f(\varepsilon) = E_0 + O(\mu^2). \tag{22}$$

Quadratic dependence on error parameter μ is an advantageous property, which should preferably be conserved when making approximations. In comparison, Weinstein's bound[2] depends on μ in a linear fashion,[35] while Temple's bound[1] is quadratic in μ, similarly to the bracketing function.

A key quantity of our approximation is $G^0|\varphi\rangle$ which we now examine in detail. Starting from

$$G^0|\varphi\rangle = |\varphi\rangle \eta^{-1} - \sum_{i=1}^{N} \frac{H_{\widetilde{i}0}}{H_{\widetilde{i}i} - \varepsilon} |\varphi_i\rangle \eta^{-1}, \tag{23}$$

we wish to see the leading terms in increasing orders of μ. To proceed, μ-dependence of basis functions is revealed first.

Introducing the notations $\langle \xi_i | \psi_0 \rangle = z_i$ and $\langle \xi_i | \chi \rangle = w_i$, components x_i of the reference can be written as

$$x_i = \mu_0 z_i + \mu w_i, \quad i = 0,\dots,N, \tag{24}$$

yielding

$$|\varphi_i\rangle = |\xi_i\rangle - |\psi_0\rangle z_i - \mu \mu_0 \big(|\psi_0\rangle w_i + |\chi\rangle z_i \big) + O(\mu^2), \quad i = 1,\dots,N. \tag{25}$$

Leading terms of the reciprocal vectors to these excited functions take the form

$$\langle \widetilde{\varphi}_i | = \langle \xi_i | - \frac{z_i}{z_0} \langle \xi_0 | + \frac{\mu}{z_0 \mu_0} \left(\frac{z_i}{z_0} w_0 - w_i \right) \langle \xi_0 | + O(\mu^2), \quad i = 1,\dots,N. \tag{26}$$

Matrix elements needed for Eq. (23) can now be expressed with μ as

$$H_{00} = E_0 + O(\mu^2), \tag{27}$$

$$H_{0i} = \mu \left(\langle \chi | H | \xi_i \rangle - E_0 w_i \right) + O(\mu^2) = \mu\, h_i + O(\mu^2), \tag{28}$$

$$H_{\widetilde{i}0} = \mu \left(\langle \xi_i | H | \chi \rangle - E_0 w_i - \frac{z_i}{z_0} \left(\langle \xi_0 | H | \chi \rangle - E_0 w_0 \right) \right) + O(\mu^2) = \mu\, \widetilde{h}_i + O(\mu^2), \tag{29}$$

$$H_{\widetilde{i}i} = \langle \xi_i | H | \xi_i \rangle - E_0 z_i^2 - \frac{z_i}{z_0} \left(\langle \xi_0 | H | \xi_i \rangle - E_0 z_i z_0 \right) + O(\mu) = p_i + O(\mu), \tag{30}$$

with h_i, $\widetilde{h}_i$, and p_i introduced for brevity. The above give rise to

$$\eta^{-1} = (E_0 - \varepsilon)^{-1} + O(\mu^2) \tag{31}$$

and

$$\frac{H_{\widetilde{i}0}}{H_{\widetilde{i}i} - \varepsilon} = \mu\, \frac{\widetilde{h}_i}{p_i - \varepsilon} + O(\mu^2). \tag{32}$$

The dependence of Eq. (23) on μ is finally obtained as

$$G^0|\varphi\rangle = |\psi_0\rangle \frac{\mu_0}{E_0 - \varepsilon} + \mu \underbrace{\left(|\chi\rangle - \sum_{i=1}^{N} \left(|\xi_i\rangle - |\psi_0\rangle z_i \right) \frac{\widetilde{h}_i}{p_i - \varepsilon} \right) \frac{1}{E_0 - \varepsilon}}_{|\zeta\rangle} + O(\mu^2). \tag{33}$$

As $\langle \psi_0 | \zeta \rangle = 0$, it readily follows, that

$$\langle \varphi | G^0 | \varphi \rangle = \frac{1}{E_0 - \varepsilon} + O(\mu^2), \tag{34}$$

hence the first term of the Taylor-expansion, Eq. (3),

$$f^{[0]} = \varepsilon + \langle \varphi | G^0 | \varphi \rangle^{-1} = E_0 + O(\mu^2) \tag{35}$$

is well behaving. For the next term of the series, we need

$$VG^0 | \varphi \rangle = (H - \varepsilon)G^0 | \varphi \rangle - | \varphi \rangle$$
$$= \mu \underbrace{\left((H - \varepsilon)|\zeta\rangle - |\chi\rangle \right)}_{|\vartheta\rangle} + O(\mu^2). \tag{36}$$

Again, since $\langle \psi_0 | \vartheta \rangle = 0$, a quadratic dependence is revealed in μ,

$$\langle \varphi | G^0 V G^0 | \varphi \rangle \sim \mu^2, \tag{37}$$

leading to

$$f^{[1]} = \varepsilon + \langle \varphi | G^0 - G^0 V G^0 | \varphi \rangle^{-1} = E_0 + O(\mu^2). \tag{38}$$

Higher orders of Eq. (3) would ideally exhibit systematic increase in the powers of μ. This apparently does not occur stepping from order zero to one. It can be shown however that an order in μ is gained at the next term of Eq. (3). To see this, let us consider

$$\langle \varphi | G^0 V = \mu \underbrace{\left(\langle \widetilde{\zeta} | (H - \varepsilon) - \langle \chi | \right)}_{\langle \widetilde{\vartheta} |} + O(\mu^2), \tag{39}$$

with

$$\langle \widetilde{\zeta} | = \left(\langle \chi | - \sum_{i=1}^{N} \left(\langle \xi_i | - z_i z_0^{-1} \langle \xi_0 | \right) h_i (p_i - \varepsilon)^{-1} \right) (E_0 - \varepsilon)^{-1}.$$

As

$$G^0 = \frac{1}{E_0 - \varepsilon} | \psi_0 \rangle \langle \psi_0 | + O(\mu), \tag{40}$$

and $\langle \widetilde{\vartheta} | \psi_0 \rangle = 0$, direct substitution of Eqs. (36), (39), and (40) results

$$\langle \varphi | G^0 V G^0 V G^0 | \varphi \rangle \sim \mu^3, \tag{41}$$

contributing a leading term proportional to μ^3 to $f^{[2]}$.

It remains to check the effect of inserting projector O_D to the right of H in Eq. (36) and to the left of H in the analogous expression behind Eq. (39). Assuming a form

$$O_D = | \psi_0 \rangle \langle \psi_0 | + O_\perp, \tag{42}$$

with $O_\perp | \psi_0 \rangle = 0$, it is easy to see that the scalings reflected by Eqs. (38) and (41) remain valid.

III. WORKING FORMULAE AND IMPLEMENTATION

We present three versions of approximation for Löwdin's bracketing function, Eq. (1). The Neumann-series often exhibited alternant behaviour in our preliminary studies, for this reason even orders are examined. Truncation beyond order zero represents the simplest approach,

$$f^{[0]} = \varepsilon + \left(x^T G^0 x \right)^{-1}. \tag{43}$$

Taking a first correction to vector x as

$$x^{(1)} = V G^0 x, \tag{44}$$

the bracketing function up to order two is calculated as

$$f^{[2]} = \varepsilon + \left(x^T G^0 x - x^T G^0 x^{(1)} + x^{(1)T} G^0 x^{(1)} \right)^{-1}. \tag{45}$$

Note that evaluation of $f^{[2]}$ requires one extra transformation by H in each step of Davidson's algorithm,[35] due to the appearance of V in $x^{(1)}$. The strictly correct form of $f^{[2]}$ would in fact necessitate two transformations by H, since $x^T G^0 V$ should figure in it instead of $x^{(1)T}$. This approximation made by Eq. (45) is not expected to be severe, in particular, the scaling behavior elaborated in Section II C persists.

Projector O_D applied once generates

$$x_D^{(1)} = (H O_D - H^0) G^0 x, \tag{46}$$

giving a straightforward nonsymmetric approximation of $f^{[2]}$ as

$$f_{OD}^{[2]} = \varepsilon + \left(x^T G^0 x - x^T G^0 x_D^{(1)} + x_D^{(1)T} G^0 x_D^{(1)} \right)^{-1}. \tag{47}$$

Expressions (43), (45), and (47) are computed at each step n of Davidson's full CI algorithm, taking the lowest eigenvector of the n-dimensional Hamiltonian matrix as reference, φ. As upper bound, $\varepsilon = c^{1T} H c^1$ is set in the first step and ε is kept fixed during the iteration. Updating at each step as, e.g., $\varepsilon = \langle \varphi | H | \varphi \rangle = H_{00}$ would be counteractive, since this results zero for η of Eq. (10) if H_{0i} and $H_{\bar{i}0}$ disappear. As a consequence, G^0 blows up if φ becomes exact. An upper bound sufficiently far from the exact eigenvalue but still in the first sector of the bracketing function[16] is hence needed for numerical stability. We finally note on ε that above choice combined with the one-dimensional projector $O_D = c^1 c^{1T}$ makes $f_{OD}^{[2]}$ divergent at the first iteration step. To avoid this problem, ε can be altered, but for simplicity we do not report values for Eq. (47) at $n = 1$.

Working formulae (43), (45), and (47) can be easily computed by any direct CI algorithm, supplementing with two subroutines to calculate the effect of G^0 according to Eq. (15) and perform the transformation by $H O_D$ based on Eq. (20). Our implementation of $f^{[0]}$ built into the code of Knowles and Handy[36] altogether takes 17 extra scalar products. Declaration of extra full CI long vectors is avoided, at the expense of disk reading or writing 32 times. Calculation of $f_{OD}^{[2]}$ takes in addition $10 + 3n$ scalar products and $16 + 3n$ disk input–output (I/O), n denoting the iteration step.

IV. ILLUSTRATIVE APPLICATIONS

A. Total energies

Accuracy of error estimates is illustrated on the example of the H_2O molecule. In Tables I-III, legitimate error of the energy and wavefunction (collected in the last two columns) is displayed together with energy error estimates. Comparing error estimates based on Weinstein's bound (3rd column) to the obvious measure of expectation value decrease (2nd column), Weinstein's bound is found orders of magnitude looser. In accordance with its linear dependence on the error

TABLE I. Energy error estimates on the example of the H_2O molecule in 6-31G* basis set, at equilibrium geometry ($R_e = 0.947\,296$ Å, $\alpha(HOH) = 105.5025°$). Start vector of the full CI iteration is an Antisymmetrized Product of Strongly Orthogonal Geminals (APSG). Iteration step is denoted by n, and expectation value at step n is designated as $E_U(n)$. Approximate lower bounds $f^{[0]}$, $f^{[2]}_{OD}$, and $f^{[2]}$ are calculated according to Eqs. (43), (47), and (45), respectively. Asterisk indicates violation of lower bound property of the value of f implied. According to step IIIC. of Ref. 35, Weinstein's bound (to the *matrix* eigenvalue) at step n is computed as $E_{\text{Weinstein}} = E_U(n) - \|r(n)\|$, $r(n)$ standing for the residual vector. Legitimate error of the energy and wavefunction at step n is given in the last two columns, for reference. Energy unit E_h is adopted. See text for further clarification.

n	$E_U(n)$ $-E_U(n-1)$	$E_U(n)$ $-E_{\text{Weinstein}}$	$E_U(n)$ $-f^{[0]}$	$E_U(n)$ $-f^{[2]}_{OD}$	$E_U(n)$ $-f^{[2]}$	$E_U(n)$ $-E_{\text{full CI}}$	$\|\varphi(n)$ $-\psi_{\text{full CI}}\|$
2	-1.42×10^{-1}	2.93×10^{-1}	1.89×10^{-2}	3.72×10^{-2}	1.97×10^{-2}	1.52×10^{-2}	6.24×10^{-2}
3	-1.36×10^{-2}	9.34×10^{-2}	1.65×10^{-3}	3.19×10^{-3}	1.82×10^{-3}	1.54×10^{-3}	2.15×10^{-2}
4	-1.36×10^{-3}	3.00×10^{-2}	$1.72 \times 10^{-4*}$	3.42×10^{-4}	2.00×10^{-4}	1.76×10^{-4}	8.46×10^{-3}
5	-1.49×10^{-4}	1.10×10^{-2}	$2.43 \times 10^{-5*}$	4.92×10^{-5}	3.03×10^{-5}	2.72×10^{-5}	3.88×10^{-3}
6	-2.18×10^{-5}	4.58×10^{-3}	$4.42 \times 10^{-6*}$	9.23×10^{-6}	5.88×10^{-6}	5.35×10^{-6}	1.89×10^{-3}
7	-4.20×10^{-6}	2.18×10^{-3}	$9.89 \times 10^{-7*}$	2.23×10^{-6}	1.30×10^{-6}	1.15×10^{-6}	8.97×10^{-4}
8	-9.26×10^{-7}	1.04×10^{-3}	$2.08 \times 10^{-7*}$	4.52×10^{-7}	2.68×10^{-7}	2.28×10^{-7}	4.03×10^{-4}
9	-1.82×10^{-7}	4.48×10^{-4}	3.9×10^{-8}	8.5×10^{-8}	5.2×10^{-8}	4.6×10^{-8}	2.06×10^{-4}
10	-3.5×10^{-8}	2.00×10^{-4}	$8.0 \times 10^{-9*}$	1.9×10^{-8}	1.1×10^{-8}	1.1×10^{-8}	1.20×10^{-4}
11	-8.0×10^{-9}	1.04×10^{-4}	$2.0 \times 10^{-9*}$	5.0×10^{-9}	3.0×10^{-9}	3.0×10^{-9}	6.85×10^{-5}
12	-2.0×10^{-9}	5.84×10^{-5}	1.0×10^{-9}	2.0×10^{-9}	1.0×10^{-9}	1.0×10^{-9}	3.45×10^{-5}

in the wavefunction,[35] Weinstein's bound typically falls in the same order of magnitude as the error of the wavefunction. (For large n, $\|\varphi(n) - \psi_{\text{full CI}}\|$ becomes proportional to μ of Section II C.) Error estimates based on the upper bound (column 2) and the bracketing function (columns 4-6) are considerably tighter, exhibiting approximately quadratic dependence on μ for large n.

Let us compare now bracketing function based error estimate $f^{[2]}_{OD}$ (column 5) to its more rigorous counterpart based on $f^{[2]}$ (column 6). It appears from Tables I-III that projection into the Davidson subspace affects results only slightly, leaving the order of magnitude unchanged, whenever the error in the wavefunction is sufficiently small (value in last column is smaller than ca. 0.3). At contrast with this, the effect is dramatic when φ gets far for the exact wavefunction, cf. $n = 2, 3$ of Table III. Watching for numerical performance, $f^{[2]}$ is definitely more favorable than $f^{[2]}_{OD}$, based on the results collected in Tables I-III. It has to be kept in mind however that the computational cost of $f^{[2]}$ at step n is comparable to that of $f^{[2]}_{OD}$ at step $2n$. Examining results from this perspective reveals that cost per performance prefers $f^{[2]}_{OD}$ over $f^{[2]}$.

The cheapest among bracketing function based error estimates, $f^{[0]}$ performs strikingly well, the estimated error falling in the same order of magnitude as the proper error. Violation of the lower bound property (indicated by asterisk) occurs often for $f^{[0]}$, however. The cheaper one of second order estimates, $f^{[2]}_{OD}$, yields a larger error bar than $f^{[0]}$, typically

TABLE II. Same as Table I with OH bond distance $2R_e = 1.894\,592$ Å.

n	$E_U(n)$ $-E_U(n-1)$	$E_U(n)$ $-E_{\text{Weinstein}}$	$E_U(n)$ $-f^{[0]}$	$E_U(n)$ $-f^{[2]}_{OD}$	$E_U(n)$ $-f^{[2]}$	$E_U(n)$ $-E_{\text{full CI}}$	$\|\varphi(n)$ $-\psi_{\text{full CI}}\|$
2	-1.21×10^{-1}	2.53×10^{-1}	2.74×10^{-2}	5.50×10^{-2}	6.54×10^{-2}	1.76×10^{-2}	1.55×10^{-1}
3	-1.19×10^{-2}	1.33×10^{-1}	5.81×10^{-3}	1.45×10^{-2}	1.26×10^{-2}	5.69×10^{-3}	1.20×10^{-1}
4	-3.71×10^{-3}	5.64×10^{-2}	$1.56 \times 10^{-3*}$	7.57×10^{-3}	3.09×10^{-3}	1.98×10^{-3}	9.29×10^{-2}
5	-1.13×10^{-3}	3.98×10^{-2}	$7.47 \times 10^{-4*}$	3.80×10^{-3}	1.70×10^{-3}	8.48×10^{-4}	6.17×10^{-2}
6	-4.96×10^{-4}	3.04×10^{-2}	4.62×10^{-4}	1.45×10^{-3}	1.53×10^{-3}	3.52×10^{-4}	3.39×10^{-2}
7	-2.40×10^{-4}	1.91×10^{-2}	$1.03 \times 10^{-4*}$	3.29×10^{-4}	2.11×10^{-4}	1.12×10^{-4}	1.68×10^{-2}
8	-8.38×10^{-5}	9.11×10^{-3}	$2.39 \times 10^{-5*}$	6.58×10^{-5}	3.73×10^{-5}	2.85×10^{-5}	8.22×10^{-3}
9	-1.95×10^{-5}	4.98×10^{-3}	$7.31 \times 10^{-6*}$	1.94×10^{-5}	1.15×10^{-5}	9.01×10^{-6}	4.59×10^{-3}
10	-6.42×10^{-6}	2.82×10^{-3}	$1.98 \times 10^{-6*}$	5.40×10^{-6}	2.96×10^{-6}	2.59×10^{-6}	2.31×10^{-3}
11	-2.01×10^{-6}	1.49×10^{-3}	$5.39 \times 10^{-7*}$	1.27×10^{-6}	6.89×10^{-7}	5.77×10^{-7}	9.95×10^{-4}
12	-4.59×10^{-7}	6.26×10^{-4}	$8.7 \times 10^{-8*}$	2.11×10^{-7}	1.24×10^{-7}	1.18×10^{-7}	5.55×10^{-4}
13	-8.3×10^{-8}	2.79×10^{-4}	$2.2 \times 10^{-8*}$	7.1×10^{-8}	3.6×10^{-8}	3.5×10^{-8}	3.77×10^{-4}
14	-2.1×10^{-8}	1.73×10^{-4}	$8.0 \times 10^{-9*}$	2.6×10^{-8}	1.5×10^{-8}	1.4×10^{-8}	2.50×10^{-4}
15	-8.0×10^{-9}	1.14×10^{-4}	$4.0 \times 10^{-9*}$	1.2×10^{-8}	7.0×10^{-9}	6.0×10^{-9}	1.46×10^{-4}
16	-4.0×10^{-9}	7.42×10^{-5}	$1.0 \times 10^{-9*}$	4.0×10^{-9}	2.0×10^{-9}	2.0×10^{-9}	7.78×10^{-5}
17	-1.0×10^{-9}	4.04×10^{-5}	1.0×10^{-9}	1.0×10^{-9}	1.0×10^{-9}	1.0×10^{-9}	4.12×10^{-5}

TABLE III. Same as Table I with OH bond distance $4R_e = 3.789\,184$ Å.

n	$E_U(n)$ $- E_U(n-1)$	$E_U(n)$ $- E_{Weinstein}$	$E_U(n)$ $- f^{[0]}$	$E_U(n)$ $- f^{[2]}_{OD}$	$E_U(n)$ $- f^{[2]}$	$E_U(n)$ $- E_{full\,CI}$	$\|\varphi(n)$ $- \psi_{full\,CI}\|$
2	-1.20×10^{-1}	1.84×10^{-1}	$1.79 \times 10^{-2*}$	$-3.06 \times 10^{-1*}$	2.11×10^{-1}	2.01×10^{-2}	4.36×10^{-1}
3	-8.14×10^{-3}	7.87×10^{-2}	1.80×10^{-2}	$-1.65 \times 10^{-1*}$	3.01×10^{-1}	1.20×10^{-2}	3.81×10^{-1}
4	-5.95×10^{-3}	1.23×10^{-1}	$5.39 \times 10^{-3*}$	4.49×10^{-2}	1.98×10^{-2}	6.02×10^{-3}	2.15×10^{-1}
5	-4.27×10^{-3}	7.85×10^{-2}	3.98×10^{-3}	5.53×10^{-2}	6.20×10^{-2}	1.75×10^{-3}	6.91×10^{-2}
6	-1.24×10^{-3}	4.71×10^{-2}	6.38×10^{-4}	2.42×10^{-3}	1.44×10^{-3}	5.04×10^{-4}	2.51×10^{-2}
7	-4.09×10^{-4}	2.05×10^{-2}	9.85×10^{-5}	2.44×10^{-4}	2.67×10^{-4}	9.48×10^{-5}	8.03×10^{-3}
8	-8.41×10^{-5}	7.30×10^{-3}	1.39×10^{-5}	3.44×10^{-5}	4.33×10^{-5}	1.06×10^{-5}	2.39×10^{-3}
9	-7.25×10^{-6}	3.90×10^{-3}	5.67×10^{-6}	2.02×10^{-5}	2.65×10^{-5}	3.38×10^{-6}	1.34×10^{-3}
10	-2.91×10^{-6}	1.52×10^{-3}	4.74×10^{-7}	1.25×10^{-6}	7.61×10^{-7}	4.61×10^{-7}	4.85×10^{-4}
11	-4.11×10^{-7}	5.16×10^{-4}	6.0×10^{-8}	1.39×10^{-7}	9.7×10^{-8}	5.0×10^{-8}	1.57×10^{-4}
12	-4.2×10^{-8}	1.92×10^{-4}	9.0×10^{-9}	4.1×10^{-8}	3.7×10^{-8}	8.0×10^{-9}	8.05×10^{-5}
13	-7.0×10^{-9}	6.64×10^{-5}	1.0×10^{-9}	6.0×10^{-9}	4.0×10^{-9}	1.0×10^{-9}	3.86×10^{-5}

1–6 times the legitimate error. In return, it represents a more reliable lower bound than $f^{[0]}$: violation of $f^{[2]}_{OD}$ appears in parallel with the deterioration of the reference function, as reflected by the last column of Table III. The reason behind
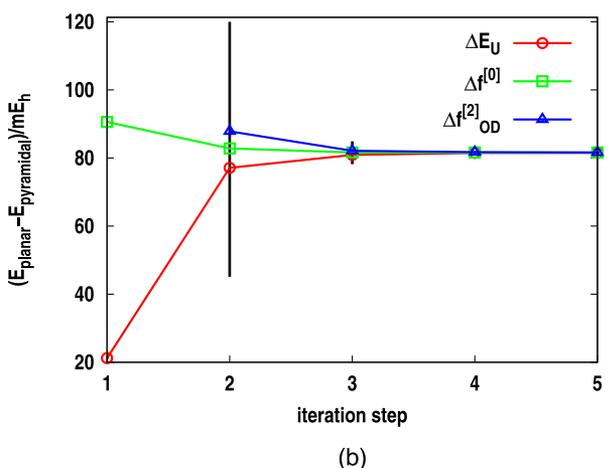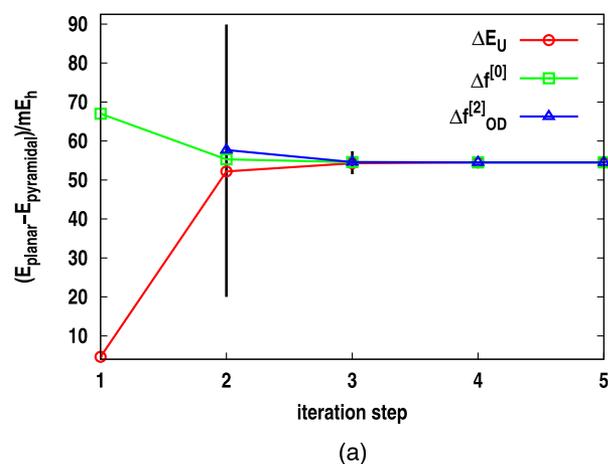

(a)

(b)

FIG. 2. Inversion barrier of the NH$_3$ molecule calculated in (a) 6-31G* and (b) 6-311G* basis sets, with core electrons frozen. Geometry is optimized at the SCF level in both cases. The Hartree-Fock function serves as start vector of the full CI iteration. Expectation value is designated as E_U. Approximate lower bounds $f^{[0]}$ and $f^{[2]}_{OD}$ are calculated according to Eqs. (43) and (47), respectively. Vertical black lines indicate error bars for the barrier, derived from E_U and $f^{[2]}_{OD}$. See text for further details.

this breakdown of the reference function upon stretching both OH bonds has been investigated recently.[37] Failure of $f^{[2]}_{OD}$ to represent a lower bound is however beneficially obvious from the reversed order of E_U and $f^{[2]}_{OD}$ (cf. the negative sign for $n = 2, 3$ in Table III). This is in contrast with $f^{[0]}$, where lower bound violation would remain hidden in lack of the legitimate error. To complete the picture, we note here that $f^{[2]}$ may also fail to represent a lower bound. This was observed at the first iteration step, both OH bonds elongated to $4R_e$. Similarly to $f^{[2]}_{OD}$, shooting above the expectation value is obtained ($E_U - f^{[2]} = -0.393$) in this case, accompanied by a large error in the reference ($\|\varphi - \psi_{full\,CI}\| = 0.545$). Breakdown of $f^{[2]}$ in such circumstances is justifiable regarding that the zero-order of Eq. (4) gets worse along with φ getting worse. We plan to explore in follow-up studies whether wavefunction error estimates extractable during iteration can be linked with the reliability of $f^{[2]}$ or $f^{[2]}_{OD}$ as lower bounds.

B. Energy barriers

Barriers of inversion are displayed for the NH$_3$ molecule in two basis sets in Fig. 2. While the barrier itself changes significantly with the basis set at this level, lower bound derived barriers behave similarly. Barrier of $f^{[0]}$ outperforms both E_U and $f^{[2]}$. It is however $f^{[2]}_{OD}$ that can be utilized (starting at step 2) to generate an error bar for the barrier by subtracting bounds on the planar and pyramidal ground state energy. An upper bound is provided by E_U(planar) $- f^{[2]}_{OD}$(pyramidal), and a lower bound is given by $f^{[2]}_{OD}$ (planar)$-E_U$ (pyramidal). As Fig. 2 shows the thus derived error bar is yet too loose at step 2, but it tightens rapidly as the iteration proceeds. In 6-31G* basis set, bounds for the barrier in mE$_h$ are $54.2 < \Delta E < 54.8$ at iteration step 4 and $54.45 < \Delta E < 54.5$ at step 5. The same values in 6-311G* basis, again in mE$_h$, are $81.2 < \Delta E < 81.9$ at step 4 and $81.5 < \Delta E < 81.6$ at step 5. These error estimates fall in the same order of magnitude as the error itself of the barrier computed as E_U(planar) $- E_U$(pyramidal), the actual values being 0.2 mE$_h$ at step 4 and 0.03 mE$_h$ at step 5 in 6-31G* basis and 0.1 mE$_h$ at steps 4 and 5 in 6-311G* basis.

V. CONCLUSION

Two simple approximations to Löwdin's bracketing function provide tight error estimation (cf. Eq. (43)) and a fairly reliable lower bound (cf. Eq. (47)) on the ground state energy during direct CI iteration, at negligible cost. Though strict lower bound property is not guaranteed, violation for Eq. (47) appears to be accompanied by a relatively large error of the reference function. Further investigations are desirable to collect numerical experience on the performance of the present schemes. Efforts towards finding tighter, preferably strict lower bound approximations may once help to decide on breaking off a lengthy iteration procedure once the error bar falls below the desired accuracy.

Approximate evaluation of Löwdin's $f(\varepsilon)$ allows future exploration of wavefunction improvement based on the variation principle for the bracketing function.[21] Second order dependence of the bracketing function on the error in the reference function can be expected to offer some advantage over variance minimization, explored previously in connection with solving the eigenvalue problem of large matrices.[38]

[1]G. Temple, Proc. R. Soc. London, Ser. A **119**, 276 (1928).
[2]D. H. Weinstein, Proc. Natl. Acad. Sci. U. S. A. **20**, 529 (1934).
[3]A. F. Stevenson, Phys. Rev. **53**, 199 (1938).
[4]A. F. Stevenson and M. F. Crawford, Phys. Rev. **54**, 375 (1938).
[5]H. Kleindienst and W. Müller, Theor. Chem. Acc. **56**, 183 (1980).
[6]N. W. Bazley, Phys. Rev. **120**, 144 (1960).
[7]N. W. Bazley and D. W. Fox, Phys. Rev. **124**, 483 (1961).
[8]P. S. C. Wang, J. Chem. Phys. **48**, 4131–4133 (1968).
[9]M. G. Marmorino, J. Math. Chem. **49**, 1535–1543 (2011).
[10]S. S. Iyengar, G. E. Scuseria, and A. Savin, Int. J. Quantum Chem. **79**, 222–234 (2000).
[11]I. Porras, D. M. Feldmann, and F. W. King, Int. J. Quantum Chem. **71**, 455–463 (1999).
[12]A. M. K. Müller, Phys. Lett. A **73**, 149–152 (1979).
[13]R. S. Varga, *Gershgorin and His Circles*, Springer Series in Computational Mathematics Vol. 36 (Springer-Verlag, Berlin, Heidelberg, 2004).
[14]P.-O. Löwdin and T.-K. Lim, Int. J. Quantum Chem. **4**(S3B), 697–702 (1969).
[15]D. A. Mazziotti, Chem. Rev. **112**, 244–262 (2012).
[16]P. O. Löwdin, Phys. Rev. A **139**, 357–372 (1965).
[17]J. Čížek and E. R. Vrscay, Int. J. Quantum Chem. **28**, 665–686 (1985).
[18]P. O. Löwdin, Int. J. Quantum Chem. **21**, 69 (1982).
[19]P. Piecuch, S. Zarrabian, J. Paldus, and J. Čížek, Phys. Rev. A **42**, 5155 (1990).
[20]H. Meißner, J. Chem. Phys. **119**, 4126 (2003).
[21]Á. Szabados and Z. Tóth, J. Math. Chem. **52**(8), 2210–2221 (2014).
[22]A. Scrinzi, Phys. Rev. A **45**, 7787 (1992).
[23]M. G. Marmorino, J. Math. Chem. **31**, 197 (2002).
[24]B. O. Roos and P. E. M. Siegbahn, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer (Plenum, New York, 1977), Vol. 3, Chap. 7.
[25]F. Riesz and B. Szőkefalvi-Nagy, *Functional Analysis* (Blackie & Son, London and Glasgow, 1956).
[26]P. R. Surján and Á. Szabados, Collect. Czech. Chem. Commun. **69**, 105 (2004).
[27]Z. Rolik, Á. Szabados, and P. R. Surján, J. Chem. Phys. **119**, 1922 (2003).
[28]P. R. Surján, Z. Rolik, Á. Szabados, and D. Kőhalmi, Ann. Phys. **13**, 223–231 (2004).
[29]I. Mayer, Theor. Chim. Acta **104**, 163–166 (2000).
[30]P. R. Nagy, P. R. Surján, and Á. Szabados, Theor. Chim. Acta **131**, 1109 (2012).
[31]Z. Gershgorn and I. Shavitt, Int. J. Quantum Chem. **2**, 751–759 (1968).
[32]P. S. Epstein, Phys. Rev. **28**, 695 (1926).
[33]R. K. Nesbet, Proc. R. Soc. London, Ser. A **A230**, 312 (1955).
[34]P. R. Surján and Á. Szabados, in *Fundamental World of Quantum Chemistry, A Tribute to the Memory of Per-Olov Löwdin*, edited by E. J. Brändas and E. S. Kryachko (Kluwer, Dordrecht, 2004), Vol. III, pp. 129–185.
[35]E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).
[36]P. J. Knowles and N. C. Handy, Comput. Phys. Commun. **54**, 75 (1989).
[37]P. Jeszenszki, V. Rassolov, P. R. Surján, and Á. Szabados, Mol. Phys. **113**, 249–259 (2014).
[38]C. W. Murray, S. C. Racine, and E. R. Davidson, J. Comput. Phys. **103**, 382–389 (1992).
[39]D. Mukherjee, Chem. Phys. Lett. **79**, 559–562 (1981).