Energy Gaps and Cohesive Energy of Ge from the Optimized Effective Potential

D. M. Bylander and Leonard Kleinman
Department of Physics, University of Texas, Austin, Texas 78712

(Received 1 July 1994)

The optimized effective potential (OEP) yields eigenfunctions which minimize the expectation value of the Hartree-Fock Hamiltonian (in this case with an added local-density-approximation correlation energy functional). Using an approximate OEP, which yields Ge eigenvalues and total energies to within a few meV of the exact OEP, we find the Ge indirect gap close to the experimental value and the direct gap somewhat too *large*. The cohesive energy is 1.2 eV less than the experimental value from which we conclude that standard correlation energy density functionals are inadequate for covalently bonded systems.

PACS numbers: 71.25.Cx, 71.25.Rk, 71.45.Nt

It is well known that Kohn-Sham (KS) [1] eigenvalues do not represent excitation energies except that the highest occupied eigenvalue equals the ionization energy [2]. Therefore the reason for the exceptionally small KS gaps in semiconductors (over 100% too small in Ge) has been a question of some concern. Using an exactly solvable model, Gunnarsson and Schönhammer [3] concluded that most of the gap error is due to the use of the local density approximation [1]. However, Godby, Schlüter, and Sham [4] using an exact relationship between the exact KS exchange-correlation potential V_{xc} and the self-energy concluded that more than 80% of the gap error (in Si) is due to an intrinsic discontinuity between the N and $N + 1$ electron V_{xc}. This discontinuity arises from a discontinuity in the slope of the xc energy functional $E_{\mathrm{xc}}[\rho]$ at $n = N$, where the number of electrons n is considered a continuous variable [5]. Because no attempt [6,7] to construct a legitimate [8] KS V_{xc} beyond the local density approximation (LDA) has resulted in much better energy gaps, the latter conclusion has become generally accepted.

We [9] recently showed that a Slater [10] exchange potential, which we called the average Fock approximation (AFA), is an approximate KS exchange potential in that it can be obtained as a *partial* functional derivative of the Fock exchange energy. We [11] then performed calculations for Ge and GaAs in which the valence electrons experienced Hedin-Lundqvist (HL) [12] correlation among themselves in addition to the AFA exchange potential, obtaining the direct and indirect gap in excellent agreement with experiment. Although this result casts some doubt on the conclusion that the KS gap errors are due mainly to the V_{xc} discontinuity, it is not conclusive because the AFA is not a true KS potential.

The optimized effective potential (OEP) is that multiplicative potential whose eigenfunctions minimize the expectation value of the Hartree-Fock (HF) Hamiltonian. The integral equation determining the OEP was first obtained by Sharp and Horton [13], who suggested two approximations. The more severe of these approximations to the integral equation resulted in the AFA. Talman and Shadwick [14] were the first to apply the OEP to atomic calculations. Krieger, Li, and Iafrate [15] cast the less severe approximation of Ref. [13] in a more transparent form (hereafter called the KLI approximation) and then [16] compared the total energy of every atom from Li to Hg in the HF, OEP, and KLI approximations. So far as we are aware neither the OEP nor KLI approximations have ever been applied to calculate the total energy or electronic structure of a solid [8].

The KLI approximation may be obtained via a simple plausibility argument without any reference to integral equations or Green's functions. Slater obtained the AFA by taking an average of the HF potential seen by the ith electron with spin σ,

$$v_{\mathrm{HF}}^{i\sigma} = \sum_{j}^{\mathrm{occ}} \int d\mathbf{r}'\, \psi_{j\sigma}^{*}(\mathbf{r}')\psi_{i\sigma}(\mathbf{r}')|\mathbf{r} - \mathbf{r}'|^{-1}\psi_{j\sigma}(\mathbf{r})/\psi_{i\sigma}(\mathbf{r}) \tag{1}$$

weighted by $\psi_{i\sigma}^{*}(\mathbf{r})\psi_{i\sigma}(\mathbf{r})/\rho_{\sigma}(\mathbf{r})$ to obtain

$$\begin{aligned} V_{\mathrm{AFA}}^{\sigma} = {} & \sum_{i,j}^{\mathrm{occ}} \int d\mathbf{r}'\, \psi_{\sigma j}^{*}(\mathbf{r}')\psi_{i\sigma}(\mathbf{r}') \\ & \times \frac{1}{|\mathbf{r} - \mathbf{r}'|}\psi_{jr}(\mathbf{r})\psi_{i\sigma}^{*}(\mathbf{r})/\rho_{\sigma}(\mathbf{r}) . \end{aligned} \tag{2}$$

Obviously, the N lowest eigenfunctions of the HF potentials of Eq. (1) yield the lowest total energy of the N-electron HF Hamiltonian and the N lowest eigenfunctions of V_{AFA} yield a somewhat higher expectation value of the HF Hamiltonian. Now, again obviously, replacing $v_{\mathrm{HF}}^{i}(\mathbf{r})$ by $v_{\mathrm{HF}}^{i}(\mathbf{r}) + C_i$, where C_i is an arbitrary constant does not affect the HF eigenfunctions. Thus taking Slater's weighted average we obtain

$$V_{\mathrm{KLI}}^{\sigma}(\mathbf{r}) = V_{\mathrm{AFA}}^{\sigma}(r) + \sum_{i} C_i n_{i\sigma}(\mathbf{r})/\rho_{\sigma}(\mathbf{r}) , \tag{3}$$

where $n_{i\sigma}(\mathbf{r}) = \psi_{i\sigma}^{*}(\mathbf{r})\psi_{i\sigma}(\mathbf{r})$ and $\rho_{\sigma} = \sum_i C_i n_{i\sigma}$. We can see that the C_i associated with states that are degenerate by symmetry must be taken to be identical to maintain the symmetry of the system and that a constant added to all the C_i just adds that constant to $V_{\mathrm{KLI}}^{\sigma}$ so that the C_i

are linearly related and one must be arbitrarily chosen. In atoms it is customary, but not necessary, to choose the C_i of the highest occupied state of spin σ to be zero so that $\sum_i C_i n_{i\sigma}(\mathbf{r})/\rho_\sigma(\mathbf{r})$ will go to zero as $\mathbf{r} \to \infty$. Our aim is to find C_i such that occupied eigenfunctions of $V^\sigma_{\mathrm{KLI}}(\mathbf{r})$ will be as similar as possible to the eigenfunctions of $v^{i\sigma}_{\mathrm{HF}} + C_i$. A plausible way of doing this is to require that $\langle \psi_{i\sigma} | V^\sigma_{\mathrm{KLI}} | \psi_{i\sigma} \rangle = \langle \psi_{i\sigma} | V^\sigma_{\mathrm{HF}} + C_i | \psi_{i\sigma} \rangle$ or in briefer notation $\overline{V}^\sigma_{\mathrm{KLI}} = \overline{v}^{i\sigma}_{\mathrm{HF}} + C_i$. Inserting C_i back into (3) yields

$$V^\sigma_{\mathrm{KLI}}(\mathbf{r}) = V^\sigma_{\mathrm{AFA}}(\mathbf{r}) + \sum_{i=1}^{N_\sigma - 1} (\overline{V}^{i\sigma}_{\mathrm{KLI}} - \overline{v}^{i\sigma}_{\mathrm{HF}}) \hat{n}_{i\sigma}(\mathbf{r})/\rho_\sigma(\mathbf{r}) \,. \tag{4}$$

Here $\hat{n}_{i\sigma}(\mathbf{r})$ is the charge density of a set of degenerate states and N_σ is the number of such occupied sets. Fractional occupation of a degenerate set does not occur in a semiconductor and presents no problem for the atom [17].

Equation (4) may be solved for $\overline{V}^\sigma_{\mathrm{KLI}}$, yielding

$$\overline{V}^{i\sigma}_{\mathrm{KLI}} = \overline{v}^{i\sigma}_{\mathrm{HF}} + \sum_{j=1}^{N_\sigma - 1} [(I - M^\sigma)^{-1}]_{ij} (\overline{V}^{j\sigma}_{\mathrm{AFA}} - \overline{v}^{\sigma j}_{\mathrm{HF}}) \,, \tag{5}$$

where I is the identity matrix and

$$M^\sigma_{ij} = \int d\mathbf{r}\, \hat{n}_{i\sigma}(\mathbf{r}) \hat{n}_{j\sigma}(\mathbf{r})/\rho_\sigma(\mathbf{r}) \,. \tag{6}$$

This result is identical to that obtained [13,15] by replacing $(\epsilon_j - \epsilon_i)^{-1}$, appearing in a Green's function of the exact OEP by $(\overline{\epsilon}_i - \epsilon_i)^{-1}$.

We first demonstrate that V^σ_{OEP} is a KS potential [18]. We then argue that V^σ_{KLI} yields results sufficiently close to those of V^σ_{OEP} that we can draw conclusions concerning correlation functionals and true KS potentials from KLI calculations. We then describe our calculations for Ge and present the results. Finally, we state our conclusions which some may find surprising.

There are two ways to separate $E_{\mathrm{xc}}[\rho]$ into $E_x[\rho] + E_c[\rho]$. In one, $E_x[\rho]$ yields the exchange energy of KS wave functions. In the other, when E_c is dropped, E_x is such that the KS equation reproduces the HF charge density and total energy. In either case E_c is defined so that $E_x + E_c = E_{\mathrm{xc}}$. Now $V_{\mathrm{xc}} = \delta E_{\mathrm{xc}}/\delta\rho$ yields that charge density which minimizes the total energy. If we assume $E_c[\rho]$ to be known to some approximation and $E_x[\rho]$ to be unknown but to represent the exchange energy of KS wave functions, then the OEP yields eigenfunctions (and therefore charge density) which minimize the total energy [19]. Therefore the OEP is the KS potential in this case and without knowledge of $E_x[\rho]$ the total energy may be obtained from $E_c[\rho]$ and the Fock operator. The HF energy of atomic Ge was calculated [16] to be −4150.7206 Ry, with the OEP energy 0.0240 Ry above the HF and KLI energy 0.0060 Ry above the OEP. The 0.0006% discrepancy between the HF and OEP energies would be included in E_c if correlation had been included in the calculation [20]. Furthermore, the OEP energy being exact except for errors introduced by the approximation to E_c, the KLI error is only 0.000 15% in this case. KLI [16] also compare their $4s_\downarrow$ and $4p_\uparrow$ eigenvalues of 0.9573 and 0.5706 Ry with the OEP values, 0.9577 and 0.5734 Ry. In these calculations valence electrons see the same exchange potential as core electrons so this KLI approximation to the OEP is much more severe than in our pseudopotential calculations where only the valence electrons have their exchange potentials averaged.

To obtain the KLI pseudopotential, a HF Dirac calculation was performed for Ge in the spin unpolarized $4s^2_{1/2}4p^2_{1/2}$ ground state. Then with the core held rigid the $4s_{1/2}4p_{1/2}$ states were self-consistently recalculated [21] using $V_{\mathrm{KLI}}(\mathbf{r})$ and valence HL correlation [12]. The unoccupied $4p_{3/2}$, $4d_{3/2}$, and $4d_{5/2}$ eigenfunctions were then calculated and the wave functions were extended nodelessly (as in Ref. [11]) from r_c back into the core to form pseudofunctions; then the Schrödinger equation was inverted to obtain pseudopotentials. The $2j + 1$ average of the $p_{1/2}(d_{3/2})$ and $p_{3/2}(d_{5/2})$ pseudopotentials was taken to obtain spin-orbit averaged pseudopotentials and the p and d pseudofunctions recalculated to construct factorized ionic pseudopotentials [22]. These pseudopotentials were used in the crystal as well as to calculate the total energy of the spin polarized KLI-HL pseudoatom [23]. The energy of the pseudoatom lies 22 meV below the −7.67 12 Ry of the AFA atom [11]. Even if the KLI percentage error (relative to the exact OEP result) were an order of magnitude larger than for the full atom, the OEP result would still lie only 1.5 meV below the KLI.

The calculation of V_{AFA} in the crystal has previously been described [11]. We use the same 10 **k**-point sample of the irreducible wedge of the Brillouin zone (BZ). There are four occupied bands except for the $(\frac{1}{8}\frac{1}{8}\frac{1}{8})$ and $(\frac{3}{8}\frac{3}{8}\frac{3}{8})$ **k** points where the highest band is twofold degenerate; thus there are 38 different $\hat{n}_{i\sigma}$, the highest lying one of which is discarded, making M_{ij} a 37 × 37 matrix. Each $\hat{n}_{i\sigma}$ is the charge from the full star of the wave vector at each **k** point and band sampled. Thus 240 different states [24] contribute to the 37 $\hat{n}_{i\sigma}$. Starting from the AFA eigenfunctions it took four iterations to converge the total energy to 0.01 meV and the individual eigenvalues to 1 meV. The most time consuming part of the calculation is updating V^σ_{AFA} after each iteration [25].

In Table I are listed the experimental and KLI direct gap at Γ, indirect gaps to X and L, and cohesive energy along with results from Ref. [11] using the core corrected LDA, the HF core-LDA valence, and AFA exchange potentials [26]. We note that the KLI gaps differ by small amounts from the AFA in the direction to make the agreement with experiment worse for each. Nevertheless, the agreement with experiment is far superior to that obtained from the LDA. Note in particular that the huge LDA $\Gamma_{2'}$ error (over 100% if spin orbit splitting were included) actually changes sign in the KLI approximation. Thus we state our first conclusion which must be somewhat

TABLE I. Energy gaps and cohesive energy (in eV) calculated in various approximations and compared with experiment.

	L_1	X_1	$\Gamma_{2'}$	$E_{\rm coh}$
LDA*	0.115	0.630	0.018	4.472
HF-LDA	0.401	0.597	0.725	4.387
AFA	0.821	1.033	1.188	2.625
KLI	0.796	0.933	1.239	2.651
Expt.	0.84	1.3 ±0.2	1.00	3.85

tentative because of our use of an approximate correlation functional: *The discrepancies between Ge energy gaps that would be obtained from the exact KS potential and experiment are small (compared to LDA discrepancies) and of random sign.* If $E_{\rm xc}$ is taken to be a functional of $\rho(\mathbf{r})$ and n, the number of electrons, then $V_{\rm xc}$ is discontinuous at integer values of n. This discontinuity should yield the lowest energy gap correctly [2]. However, our KLI potentials assume a fixed number of electrons, have no discontinuities, and therefore no eigenvalue other than the highest occupied represents an excitation energy. Thus there is no *a priori* reason to expect its better $V_{\rm xc}$ (in the sense that it yields a better total energy) to yield better energy gaps. Because of the similarities among diamond structure semiconductors, we would expect similar corrections in Si gaps. For reasons discussed in Ref. [11], the $\Gamma_{2'}$ level is extremely sensitive to changes in potential and L_1 somewhat less so. The bottom of the Si conduction bands lies close to X_1. We see in Table I that the KLI potential yields a much smaller improvement for this gap than for the others. Similarly the Si direct gap is between $\Gamma_{25'}$ and Γ_{15}. In Ge the LDA, KLI, and experimental values for this gap are 2.555, 2.800, and 3.25 eV, respectively. Thus we expect the KLI potential to yield only moderate improvements to the Si gaps. Therefore the importance of our first conclusion is not that we obtained very good direct and indirect gaps in a single case, but rather that KS gaps are not always much too small as had been previously believed. It is perhaps worth pointing out that the HF indirect gaps of Si and Ge are about 5 times larger than the experimental values [27]. In many-body quasiparticle calculations [28] the effect of correlation is to screen the Fock operator and thus reduce all gaps to their experimental values. In density functional theory, however, the exchange and correlation functionals (and hence potentials) are additive. Thus our KLI calculation without correlation, rather than approximating the HF gaps, will yield gaps that are smaller than experiment. Note that n may be taken to be a continuous variable in the KLI approximation and the integer discontinuity calculated for atoms [15]. It would be extremely tedious to calculate the discontinuity for a solid and furthermore the gap it would yield would be the HF gap augmented by the added correlation potential.

The KLI total energy lies 48 meV below the AFA resulting in a 26 meV increase in the cohesive energy. Since the atomic fourfold ionization energy is about 0.26 eV above the experimental value after core relaxation effects have been included [29], the 1.2 eV error in the KLI cohesive energy represents an error of about 0.94 eV in the crystal's total energy. Although the LDA cohesive energy is in better agreement with experiment, it comes from crystal and atomic energies 4.09 and 3.47 eV larger than experiment so that KLI total energies are actually quite a bit better than the LDA. Nevertheless, the fact that the KLI cohesive energy is so poor leads us to our second conclusion: *There is no simple correlation energy density functional which is adequate for both atoms and covalently bonded crystals.* Exchange depends only on ground state eigenfunctions, whereas correlation requires the admixture of excited configurations and therefore its functional dependence on the ground state charge density is much more indirect. If we assume that the entire cohesive energy error arises from correlation, we conclude that a functional which yields the correct total energy in the atom underestimates the correlation energy in each Ge covalent bond by 0.6 eV. We believe that it is a general result that for any standard correlation functional the errors in the correlation energy for two different structures will differ by a significant amount. Not only have we found this to be true for the Ge atom and crystal, but Phillips [30] found it to be the case when comparing small Si clusters with different structures. We would predict that the correlation energy of covalent bonds is generally underestimated but would not use these results to predict that the correlation energy causes an underestimate of the cohesive energy for crystals other than semiconductors

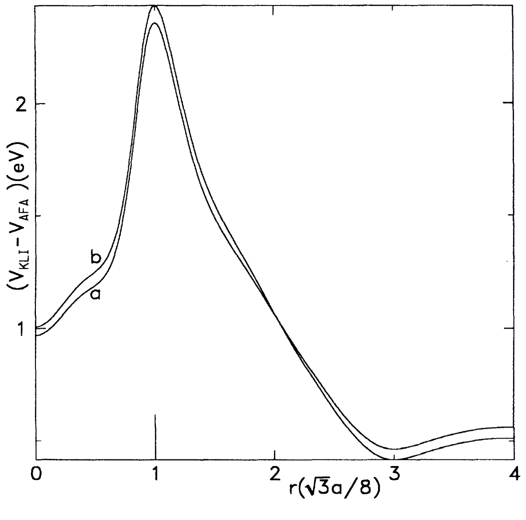

FIG. 1. The difference between the KLI and AFA potentials along the (111) direction in Ge. The atom is at $r = 1$, the middle of the covalent bond is at $r = 0$, and $r = 4$ is halfway to the far atom. In curve a, $V_{\rm AFA}$ is calculated from the same self-consistent KLI wave functions that $V_{\rm KLI}$ is; in curve b, $V_{\rm AFA}$ is calculated from its own self-consistent wave functions.

when the exchange energy is calculated nearly exactly as it is in the KLI approximation. Because the KLI approximation yields the exchange energy nearly exactly, one would expect it to always yield better total energies than the LDA; however, it is possible that a fortuitous cancellation of LDA exchange and correlation errors might in some cases make their sum less than the KLI correlation error.

Because the differences between the KLI and AFA total energy and energy gaps is small, one might conclude that their potentials differ only slightly. That this is not the case is shown in Fig. 1. Finally, we note that if it were possible to construct an accurate correlation functional of KS ground state eigenfunctions [31], the KLI approximation could be applied to both exchange and correlation and accurate results obtained for total energies.

This work was supported by the University of Texas High Performance Computing Facility, the Welch Foundation (Houston, Texas), and the NSF under Grant No. DMR 9313645.

---

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[2] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982).

[3] O. Gunnarrson and K. Schönhammer, Phys. Rev. Lett. **56**, 1968 (1986).

[4] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

[5] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[6] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **30**, 5777 (1984).

[7] Yu-Tang Shen, D. M. Bylander, and Leonard Kleinman, Phys. Rev. B **36**, 3465 (1987).

[8] Y. Li, J. B. Krieger, M. R. Norman, and G. J. Iafrate [Phys. Rev. B **44**, 10 437 (1991)] have applied the OEP to the self-interaction correction to LDA exchange and correlation. This might be considered a legitimate $V_{\rm KS}$ for an *ad hoc* Hamiltonian. By OEP in this paper we will mean OEP applied to a Hamiltonian containing Fock exchange.

[9] Leonard Kleinman, Phys. Rev. B **49**, 14 197 (1994).

[10] J. C. Slater, Phys. Rev. **81**, 385 (1951).

[11] Y. M. Gu, D. M. Bylander, and Leonard Kleinman, Phys. Rev. B **50**, 2227 (1994).

[12] L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2064 (1971).

[13] R. T. Sharp and G. K. Horton, Phys. Rev. **90**, 317 (1953).

[14] J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976).

[15] J. B. Krieger, Yan Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).

[16] Yan Li, J. B. Krieger, and G. J. Iafrate, Phys. Rev. A **47**, 165 (1993).

[17] In standard atomic calculations a strictly radial potential is obtained (which does depend on which of the states are assumed occupied in a partially filled shell). Fractional occupation can be problematical only if it is allowed to destroy the symmetry of the system. That does not happen for the atom which is always assumed spherical.

[18] V. Sahni, J. Gruenbaum, and J. P. Perdew, Phys. Rev. B **26**, 4371 (1982).

[19] The OEP is shown in Ref. [15] to be that multiplicative potential whose eigenfunctions minimize the expectation value of a Hamiltonian containing a potential operator such as Fock exchange, irrespective of whether or not the Hamiltonian contains a correlation functional. For example, Eqs. (4)–(6) for the KLI approximation are unchanged in form when correlation is included but depend on the correlation potential through the eigenfunctions which determine $\hat{n}_{j\sigma}$, $V^{\sigma}_{\rm AFA}(\mathbf{r})$, $\bar{v}^{i\sigma}_{\rm HF}$, etc.

[20] One might ask whether the LSDA to $E_c[\rho]$ should be thought of as including this. Because E_c represents the difference between the total energy and the energy obtained from a single Slater determinant whether HF or LSDA eigenfunctions are used as basis functions and because HF and LSDA eigenfunctions are identical for the free electron gas from which the LSDA is obtained, the answer falls within the uncertainty in the LSDA.

[21] Since in this case the system is spin unpolarized and each $p_{1/2}$ function contains spins of both signs, we drop the subscript and superscript σ's in Eq. (4). Note that $V_{\rm AFA}$ averages the exchange potential seen by the valence electrons due to all the electrons. In this case there is only one $\hat{n}_i(\mathbf{r})$, the charge density of the two $4s_{1/2}$ wave functions, while $\rho(\mathbf{r})$ is the charge density of the four valence electrons.

[22] Leonard Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[23] In this case $V^{\downarrow}_{\rm KLI}$ for the single down spin electron reduces to the HF potential, whereas $\Lambda^{\uparrow}_{\rm KLI}$ is again constructed from a single $\hat{n}_{i\sigma}$, that of the $4s_{\uparrow}$ function, while $\rho_{\uparrow}(\mathbf{r})$ is the charge density of $4s_{\uparrow}p^2_{\uparrow}$.

[24] There are 256 points sampled in the full Brillouin zone but the eightfold star of the twofold degenerate Λ_3 state at $(\frac{1}{8}\frac{1}{8}\frac{1}{8})$ whose $\hat{n}_{i\sigma}$ is discarded reduces this by 16.

[25] In these exploratory calculations we were interested in both AFA and KLI results but to save computational time we would recommend starting the KLI calculation from LDA eigenfunctions.

[26] As in Ref. [11] the spin orbit splitting has been averaged out of the $\Gamma_{2'}$ and L_1 experimental gaps. The negative and low lying positive gaps not listed in Table I differ from their AFA values in Ref. [11] by 0.05 eV or less.

[27] A. Svane, Phys. Rev. B **35**, 5496 (1987).

[28] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[29] In Ref. [11] we assumed that core relaxation in LDA and HF atoms is similar. Using HF core relaxation we find the fourfold ionization energy of the KLI pseudoatom is 0.26 eV above the experimental value rather than the 0.25 eV below found using LDA core relaxation.

[30] J. C. Phillips, Phys. Rev. B **47**, 14 132 (1993).

[31] M. E. Casida, Phys. Rev. A **51**, 2005 (1995).