# The Variational Method and Excited Electronic States

## Resources

1. T. Helgaker, P. Jørgensen and J. Olsen ***Molecular Electronic-Structure Theory***, Chapter 4 (Wiley, 2000).
2. A. Szabo and N. S. Ostlund, ***Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory***, Chapters 1 and 2 (Dover Publications Inc., 1989).
3. C. J. Cramer, ***Essentials of Computational Chemistry: Theories and Models***, Chapter 14 (Wiley, 2nd edition 2004).

## General formulation of the variational principle

Consider the time-independent electronic Schrödinger equation:

$$\hat{H}\left|\Phi_n\right\rangle = E_n\left|\Phi_n\right\rangle \tag{1}$$

where $\hat{H}$ is the electronic Hamiltonian with eigenstates $|\Phi_n\rangle$ and eigenvalues $E_n$. The Hamiltonian operator is Hermitian, hence, the eigenvalues $E_n$ are real and the eigenstates can be chosen to be orthonormal (see section 1.1.6 in reference 2):

$$\left\langle\Phi_n \,|\, \Phi_m\right\rangle = \delta_{nm} \tag{2}$$

The eigenstates $|\Phi_n\rangle$ represent the wave functions for the electronic states and the $E_n$ their energies. $n = 0$ corresponds to the ground state, while $n > 0$ correspond to electronic states. The energy of an approximation $\left|\tilde{\Phi}_n\right\rangle$ to an eigenstate of the Hamiltonian can be written as the expectation value:

$$E\left[\tilde{\Phi}_n\right] = \frac{\left\langle\tilde{\Phi}_n\middle|\hat{H}\middle|\tilde{\Phi}_n\right\rangle}{\left\langle\tilde{\Phi}_n\middle|\tilde{\Phi}_n\right\rangle} \tag{3}$$

where the notation $\left[\tilde{\Phi}_n\right]$ indicates that the energy is a functional of the wave function, i.e. it depends on the form of the function $\left|\tilde{\Phi}_n\right\rangle$.

The *variational principle* can be stated in the following general way: Each solution to the Schrödinger equation represents a stationary point of the energy functional $E\left[\tilde{\Phi}_n\right]$ and viceverse, i.e. *there is a one-to-one relationship between the solutions to the Schrödinger equation and the stationary points of the energy functional.*

In the first lecture of this course, we have proven that each stationary point of the energy functional is a solution to the Schrödinger equation (see notes on the variational method). Let's demonstrate here that, conversely, all solutions to the electronic Schrödinger equation represent stationary points of the energy functional. We start by considering an eigenstate of the Schrödinger equation, $|\Phi_n\rangle$, and express $\left|\tilde{\Phi}_n\right\rangle$ as $|\Phi_n\rangle$ plus an infinitesimal change:

$$\left|\tilde{\Phi}_n\right\rangle = \left|\Phi_n\right\rangle + \left|\delta\Phi_n\right\rangle \tag{4}$$

by inserting the right hand side of eq 4 in eq 3, we obtain an expression for the expectation value $E\left[\tilde{\Phi}_n\right] = E\left[\Phi_n + \delta\Phi_n\right]$ :

$$\begin{aligned} E\left[\Phi_n + \delta\Phi_n\right] &= \frac{\left\langle\Phi_n\middle|\hat{H}\middle|\Phi_n\right\rangle + \left\langle\Phi_n\middle|\hat{H}\middle|\delta\Phi_n\right\rangle + \left\langle\delta\Phi_n\middle|\hat{H}\middle|\Phi_n\right\rangle + \left\langle\delta\Phi_n\middle|\hat{H}\middle|\delta\Phi_n\right\rangle}{\left\langle\Phi_n \,|\, \Phi_n\right\rangle + \left\langle\Phi_n \,|\, \delta\Phi_n\right\rangle + \left\langle\delta\Phi_n \,|\, \Phi_n\right\rangle + \left\langle\delta\Phi_n \,|\, \delta\Phi_n\right\rangle} \\ &= \frac{E_n + \left\langle\Phi_n\middle|\hat{H}\middle|\delta\right\rangle + \left\langle\delta\Phi_n\middle|\hat{H}\middle|\Phi_n\right\rangle + \left\langle\delta\Phi_n\middle|\hat{H}\middle|\delta\Phi_n\right\rangle}{1 + \left\langle\Phi_n \,|\, \delta\Phi_n\right\rangle + \left\langle\delta\Phi_n \,|\, \Phi_n\right\rangle + \left\langle\delta\Phi_n \,|\, \delta\Phi_n\right\rangle} \end{aligned} \tag{5}$$

where the second equality follows from eq 1 and from the fact that the eigenfunctions of the Hamiltonian are normalized. In order to simplify this expression, we can use the Taylor expansion $\frac{1}{1+x} = 1 - x + O(x^2)$ with $x = \langle \Phi_n \,|\, \delta\Phi_n \rangle + \langle \delta\Phi_n \,|\, \Phi_n \rangle + \langle \delta\Phi_n \,|\, \delta\Phi_n \rangle$. By ignoring second-order terms, we obtain:

$$E\left[\Phi_n + \delta\right] = E_n + \left\langle \Phi_n \middle| \hat{H} - E_n \middle| \delta\Phi_n \right\rangle + \left\langle \delta\Phi_n \middle| \hat{H} - E_n \middle| \Phi_n \right\rangle \tag{6}$$

Since $|\Phi_n\rangle$ is an eigenstate of the Hamiltonian we can use eq 1 and:

$$\langle \Phi_n | \, \hat{H} = \langle \Phi_n | \, E_n \tag{7}$$

which finally leads to:

$$E\left[\Phi_n + \delta\right] = E_n + \langle \Phi_n \,|\, E_n - E_n \,|\, \delta\Phi_n \rangle + \left\langle \delta\Phi_n \middle| \hat{H} - \hat{H} \middle| \Phi_n \right\rangle = E_n \tag{8}$$

Therefore, the first-order variation of the energy functional vanishes for an expansion around an eigenstate of the Hamiltonian. Since the points for which the energy functional does not change to first order for an infinitesimal change $|\delta\Phi_n\rangle$ are the stationary points of the energy functional, we have demonstarted that a solution to the Schrödinger equation represents a stationary point of the energy functional.

## The variational method

The general formulation of the variational principle is the basis of the *variational method* in electronic structure theory. Given an ansatz for the electronic wave function (a form of the wave function in terms of numerical parameters), the best approximations to the electronic states and their energies within this model are obtained by variational optimization of the energy function, i.e. by finding the values of the parameters that make the energy stationary.

Note that in formulating the variational principle we have considered a generic eigenstate of the Hamiltonian. This means that the variational method can be applied in calculations of the ground state as well as the excited states. Note also that we have demonstrated the relationship of an eigenstate with a generic extremum of the energy functional, not just a minimum. Therefore, an electronic state can correspond to a saddle point or a maximum. This is the case for excited states for a linear wave function ansatz (as in the configuration interaction method), and, typically, also for a nonlinear ansatz (as in Hartree-Fock calculations), as we will shortly see.

## The linear variational problem and excited states

For a linear wave function ansatz, an approximate wave function is expressed as a linear combination of a set of $M$ $N$-electron functions:

$$\left|\tilde{\Phi}\right\rangle = \sum_{i=1}^{M} c_i \left|\Psi_i\right\rangle \tag{9}$$

For simplicity, we consider an orthonormal set of basis functions:

$$\langle \Psi_i \,|\, \Psi_j \rangle = \delta_{ij} \tag{10}$$

A linear combination for the wave function is used in the method of configuration interaction and, therefore, plays an important role in electronic structure theory. As we will see, the linear wave function ansatz leads to a simple realization of the variational method.

The energy is a function of the expansion coeffients $\{c_i\}$ as can be seen by expressing it as the expectation value:

$$E = \frac{\left\langle \tilde{\Phi} \middle| \hat{H} \middle| \tilde{\Phi} \right\rangle}{\left\langle \tilde{\Phi} \middle| \tilde{\Phi} \right\rangle} = \frac{\sum_{ij} c_i c_j \left\langle \Psi_i \middle| \hat{H} \middle| \Psi_j \right\rangle}{\sum_i c_i^2} = \frac{\sum_{ij} c_i c_j H_{ij}}{\sum_i c_i^2} \tag{11}$$

where in the last equality we have introduced the notation $H_{ij}$ for the elements of the Hamiltonian matrix in the basis $\{|\Psi_i\rangle\}$ (the configuration interaction matrix). The variational method applied to a linear wave function

ansatz consists in finding a set of coefficients for which the energy is stationary. This can be done by setting the first-order derivatives of the energy expectation value with respect to the coefficients to zero:

$$\frac{\partial E}{\partial c_i} = 0 \tag{12}$$

To find an expression for the gradient of the energy, it is convenient to first rewrite eq 11 as:

$$E \sum_i c_i^2 = \sum_{ij} c_i c_j H_{ij} \tag{13}$$

Differentiating both sides of this equation, we obtain:

$$\frac{\partial E}{\partial c_i} \sum_i c_i^2 + 2E c_i = 2 \sum_j c_j H_{ij} \tag{14}$$

where, for the right hand side, we have used that $H_{ij} = H_{ji}$, since the Hamiltonian is an Hermitian operator. Therefore, the first-order derivatives of the energy with respect to the coefficients are:

$$\frac{\partial E}{\partial c_i} = 2 \frac{\sum_j c_j H_{ij} - E c_i}{\sum_i c_i^2} \tag{15}$$

The stationarity condition, eq 12, then gives:

$$\sum_j c_j H_{ij} = E c_i \tag{16}$$

This equation can be written in matrix notion as:

$$\mathbf{Hc} = E\mathbf{c} \tag{17}$$

where $\mathbf{H}$ is the matrix with elements $H_{ij}$ and $\mathbf{c}$ is a column vector with elements $c_i$. Eq 18 is a standard eigenvalue problem, which can be solved to yield $M$ eigenvectors $\mathbf{c}^n$ and associated eigenvalues $E_0 \leq E_1 \leq \cdots . \leq E_{M-1}$ (roots of the eigenvalue problem). The eigenvectors $\mathbf{c}^n$ give the approximations to the wave functions of the electronic states:

$$\left|\tilde{\Phi}_n\right\rangle = \sum_{i=1}^{M} c_i^n \left|\Psi_i\right\rangle \tag{18}$$

and the eigenvalues their respective energies. In particular, $E_0$ is the best possible approximation to the energy of the ground state in the space spanned by the basis functions $\{|\Psi_i\rangle\}$, $E_1$ is the best possible approximation to the first excited state, and so on.

We can characterize further the solutions of the linear variational problem by analyzing the elements of the electronic Hessian (second-order derivatives of the energy with respect to the variational parameters) at these stationary points. We start by taking the second-order derivatives of both sides of eq 13:

$$\frac{\partial E}{\partial c_i \partial c_j} \sum_i c_i^2 + 2\frac{\partial E}{\partial c_i} c_j + 2\frac{\partial E}{\partial c_j} c_i + 2E\delta_{ij} = 2H_{ij} \tag{19}$$

At a stationary solution $\left|\tilde{\Phi}_k\right\rangle$, the first-order derivatives vanish, therefore:

$$\frac{\partial E_k}{\partial c_i \partial c_j} = 2(H_{ij} - E_k \delta_{ij}) \tag{20}$$

where we have further used that the eigenvectors of the Hamiltonian matrix are normalized. Since the Hessian in the basis of the eigenvectors of the Hamiltonian is given by:

$$\frac{\partial E_k}{\partial c_n \partial c_m} = 2(H_{nm} - E_k)\delta_{nm} \tag{21}$$

the eigenvalues of the electronic Hessian are:

$$\frac{\partial E_k}{\partial c_n \partial c_n} = 2(E_n - E_k) \tag{22}$$

This equation shows that the eigenvalues of the electronic Hessian at a stationary state $\left|\tilde{\Phi}_k\right\rangle$ are equal to twice the excitation energies from state $\left|\tilde{\Phi}_k\right\rangle$ to states $\left|\tilde{\Phi}_n\right\rangle$. Moreover, since $E_n < E_k$ for $n < k$, we see that the electronic Hessian has $k$ negative eigenvalues, which means that the ground state corrsponds to a minimum, *the first excited state to a first-order saddle point, the second excited state to a second-order saddle point, and so on.*

Another important property of the solutions of a linear variational problem is given by a theorem called *Cauchy's interlace theorem*. According to this theorem, given two sets of basis functions, $\{|\Psi'_i\rangle\}$ and $\{|\Psi''_i\rangle\}$, with $\{|\Psi''_i\rangle\}$ including $\{|\Psi'_i\rangle\}$, the eigenvalues of the Hamiltonian matrix in the two linear variational spaces are such that (for a proof of the theorem see, e.g., Exercise 4.2 in reference 1):

$$E''_k < E'_k \tag{23}$$

In the limit that $\{|\Psi'_i\rangle\}$ represents a *complete* space (infinite terms in the linear expansion), we obtain the result that for a finite subspace the $k$th eigenvalue of the Hamiltonian matrix is an upper bound to the energy of the $k$th exact solution to the Schrödinger equation. This implies that not only $E_0$ is an upper bound to the exact ground state energy, but also that $E_1$ is an upper bound to the exact first excited state energy, $E_2$ is an upper bound to the exact second excited state energy, and so on. The *Cauchy's interlace theorem* is illustrated in Figure 1, which shows the energies of the five lowest $^1\Sigma^+$ electronic states of the BH molecule calculated using the d-aug-cc-pVDZ atomic basis set and the configuration interaction method with increasingly large linear variational spaces, up to the space including all configurations that can be generated by excitations from the Hartree-Fock ground state determinant (full configuration interaction). We can see that as the dimension of the variational space increases, the energy of each state decreases monotonically towards the value obtained with full configuration interaction.

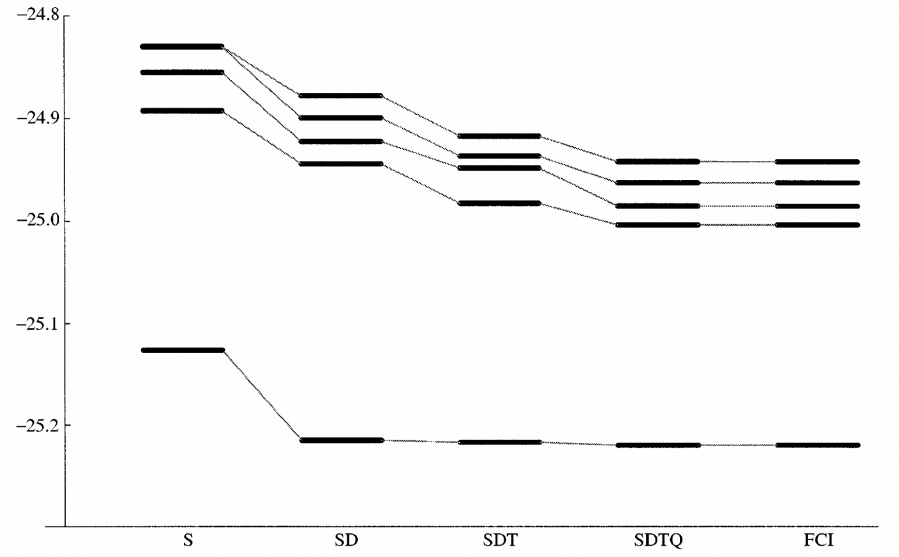


Figure 1: Energies (in Hartree) of the five lowest $^1\Sigma^+$ electronic states of the BH molecule at the ground-state equilibrium bond length calculated with the d-aug-cc-pVDZ atomic basis set and the configuration interaction method with different dimensions of the linear space formed by the configurations generated by excitation from the ground-state Hartree-Fock determinant. The full configuration interaction (FCI) calculation includes all possible configurations that can be generated within the d-aug-cc-pVDZ basis set. From Figure 4.1 of reference 1.

In summary, for a linear expansion of the wave function in a finite basis of $N$-electron functions $\{|\Psi_i\rangle\}$, the variational method is equivalent to solving an eigenvalue problem, where the solutions are the eigenvectors of the Hamiltonian matrix in the basis $\{|\Psi_i\rangle\}$, and correspond to the best approximations to the electronic states according to the variational principle. The state with lowest energy (lowest eigenvalue of the Hamiltonian matrix) is the ground state and corresponds to a minimum of the energy function, higher energy states are the

excited states, which correspond to saddle point, with the first excited state being a first-order saddle point, the second excited state a second-order saddle point, and so on. The eigenvalue of a state is an upper bound to the exact energy of that state.

## The nonlinear variational problem and excited states

For the Hartree-Fock and Khon-Sham methods, it can be shown that the form of the approximate wave function is:

$$\left|\tilde{\Phi}\right\rangle = |\Psi_0\rangle + \sum_{i,a} c_i^a \left|\Psi_i^a\right\rangle + \sum_{i<j} \sum_{a,b} c_i^a c_i^b \left|\Psi_{ij}^{ab}\right\rangle + \cdots \tag{24}$$

where $\{|\Psi_i^a\rangle\}$ and $\{\left|\Psi_{ij}^{ab}\right\rangle\}$ are singly and doubly excited Slater determinants with respect to a reference $N$-electron Slater determinant wave function $|\Psi_0\rangle$, and $c_i^a$ are the variational parameters, which represent rotation angles that mix occupied orbitals $i$ with unoccupied orbitals $a$. The form of the wave function of eq 25 is more complicated than the linear expansion of eq 9, as the variational parameters appear in a nonlinear fashion.

We can still apply the variational method to nonlinear wave functions within Hartree-Fock and Kohn-Sham calculations to obtain approximations to the ground and excited state energies. This is done by carrying out a nonlinear variational optimization (the SCF procedure) for the different states separately, and leads to solutions that do *not* provide a diagonal representation of the Hamiltonian and are *not* orthogonal to each other.

To examine the properties of Hartree-Fock nonlinearly optimized energies, in comparison to linearly optimized energies, let's consider a variationally optimized Hartree-Fock state $\left|\tilde{\Phi}\right\rangle$. The optimized $\left|\tilde{\Phi}\right\rangle$ state can be expanded in the complete set of exact orthonormal eigenvectors $|\Phi_n\rangle$ of the Hamiltonian matrix:

$$\left|\tilde{\Phi}\right\rangle = \sum_n c_n |\Phi_n\rangle \tag{25}$$

Taking $\left|\tilde{\Phi}\right\rangle$ normalized, the expectation value of the energy is:

$$E = \sum_n c_n^2 H_{nn} = \sum_n c_n^2 E_n \tag{26}$$

Using that $\sum_n c_n^2 = 1$, we can rewrite this equation into:

$$E - E_m = \sum_n c_n^2 (E_n - E_m) \tag{27}$$

from which we can see that E is equal to or greater than $E_m$, an exact eigenvalue of the Hamiltonian, only if:

$$-\sum_{n<m} c_n^2 (E_n - E_m) \leq \sum_{n>m} c_n^2 (E_n - E_m) \tag{28}$$

Since there are no states lower than the ground state, the above condition is always verified for $E_m = E_0$, the ground state energy. For another state $m$ different from the ground state, the condition is verified if $\left|\tilde{\Phi}\right\rangle$ is orthogonal to all states of energy lower than $E_m$ (since in that case $c_n$=0 for $n < m$). However, this is generally not the case in practical calculations. The implication of this is that, in general, *the energies obtained from Hartree-Fock nonlinear variational optimization of excited states while all representing upper bounds to the exact ground state energy, do not necessarily represent upper bounds to the exact excited state energies.* This is exemplified in Figure 2, which shows the energies of the $1s^2$ ground state and $1s2s$ excited state of the Helium atom calculated with the Hartree-Fock method and the t-aug-cc-pVTZ basis set as compared to the exact energies. While both nonlinearly variationally optimized energies are above the exact ground state energy, the Hartree-Fock energy of the $1s2s$ excited state is below the exact excited state energy, showing that the Hartree-Fock calculation for the excited state does not provide an upper bound to the exact excited state energy. We may use the $1s^2$ and $1s2s$ Hartree-Fock states as a basis for a configuration interaction calculation using the linear variational method seen previously. This would result in new linearly optimized energies representing upper bounds to the exact energies of the two states (see Figure 2).

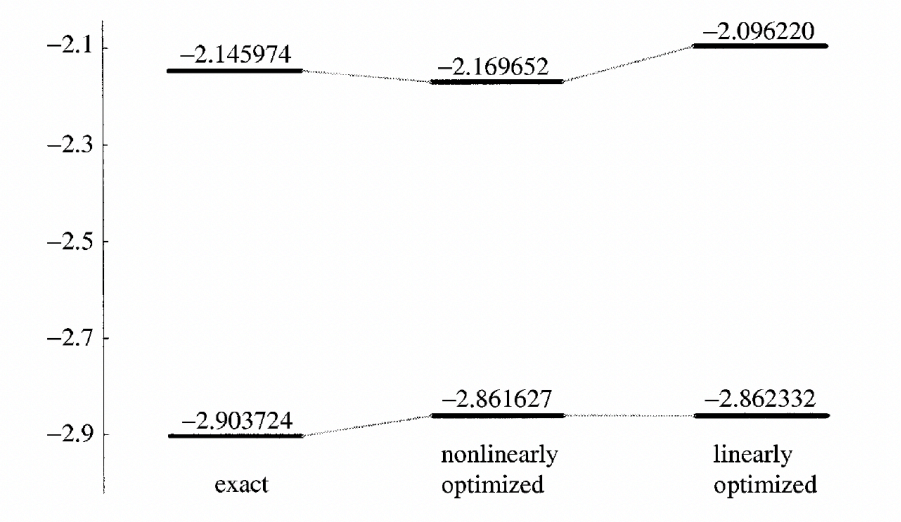


Figure 2: Approximate energies (in Hartree) of the $1s^2$ ground state and $1s2s$ excited state of the Helium atom calculated with the Hartree-Fock method and the t-aug-cc-pVTZ basis set (nonlinearly optimized) as compared to the exact energies. The energies obtained from a configuration interaction calculation in the basis of the Hartree-Fock states (linearly optimized) are also shown. From Figure 4.2 of reference 1.

In summary, the Hartree-Fock and Kohn-Sham methods use a nonlinear ansatz of the wave function. Therefore, the variational method applied within Hartree-Fock and Kohn-Sham calculations correspond to a nonlinear variational optimization, which is carried out separately for the ground and excited states, producing nonorthogonal approximations to these states. While all nonlinearly variationally optimized energies are upper bounds to the exact energy of the ground state, the approximate excited state energies are not necessarily upper bounds to the exact excited state energies. Despite this complication, the variational method can still be applied and provides useful approximations to the excited states.

## Open-shell singlet states and the sum rule

Usually, ground states are closed-shell systems. Take, for example, the ground state of the $H_2$ molecule at the equilibrium bond length. In a minimal basis, the ground state is described by the restricted closed-shell Slater determinant:

$$\left|\psi_1\overline{\psi}_1\right\rangle = \frac{1}{\sqrt{2}}\left[\psi_1(1)\psi_1(2)\right]\left(\alpha(1)\beta(2)-\alpha(2)\beta(1)\right) \tag{29}$$

where $\psi_1$ is the spatial molecular orbital obtained as a symmetric combination of two $1s$ atomic orbitals ($\sigma_g$ bonding orbital) and $(\alpha(1)\beta(2)-\alpha(2)\beta(1))$ represents the spin function for a state with singlet spin multiplicity.

On the other hand, excited states con often be open-shell systems, with multiple unpaired electrons. Consider the determinant obtained by excitation of an electron from the $\psi_1$ to the $\psi_2$ orbital of minimal basis $H_2$, where $\psi_2$ is the spatial molecular orbital obtained as an antisymmetric combination of two $1s$ atomic orbitals ($\sigma_u^*$ antibonding orbital):

$$\left|\psi_1\overline{\psi}_2\right\rangle = \frac{1}{\sqrt{2}}\left[\psi_1(1)\alpha(1)\psi_2(2)\beta(2)-\psi_1(2)\alpha(2)\psi_2(1)\beta(1)\right] \tag{30}$$

The expression for this determinant does not contain the spin function identified in eq 29, and, therefore, the determinant is not a pure singlet state. We can form a pure singlet wave function by taking the following appropriate (spin-adapted) linear combination of the two determinants $\left|\psi_1\overline{\psi}_2\right\rangle$ and $\left|\overline{\psi}_1\psi_2\right\rangle$:

$$\left|{}^1\Psi_1^2\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\psi_1\overline{\psi}_2\right\rangle + \left|\overline{\psi}_1\psi_2\right\rangle\right) = \frac{1}{2}\left[\psi_1(1)\psi_2(2)+\psi_1(2)\psi_2(1)\right]\left(\alpha(1)\beta(2)-\alpha(2)\beta(1)\right) \tag{31}$$

where we recognize the spin function for a singlet state. Smilarly, we can obtain a pure triplet wave function:

$$\left|{}^3\Psi_1^2\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\psi_1\overline{\psi}_2\right\rangle - \left|\overline{\psi}_1\psi_2\right\rangle\right) = \frac{1}{2}\left[\psi_1(1)\psi_2(2)-\psi_1(2)\psi_2(1)\right]\left(\alpha(1)\beta(2)+\alpha(2)\beta(1)\right)$$

Since:

$$\left|\psi_1\overline{\psi}_2\right\rangle = \frac{1}{\sqrt{2}}\left(\left|{}^1\Psi_1^2\right\rangle + \left|{}^3\Psi_1^2\right\rangle\right) \tag{32}$$

we obtain that the energy expectation value of the mixed-spin open-shell determinant $\left|\psi_1\overline{\psi}_2\right\rangle$ is:

$$E\left(\left|\psi_1\overline{\psi}_2\right\rangle\right) = \left\langle\psi_1\overline{\psi}_2\right|\hat{H}\left|\psi_1\overline{\psi}_2\right\rangle = \frac{1}{2}\left(\left\langle{}^1\Psi_1^2\right|\hat{H}\left|{}^1\Psi_1^2\right\rangle + \left\langle{}^3\Psi_1^2\right|\hat{H}\left|{}^3\Psi_1^2\right\rangle\right) = \frac{1}{2}\left(E\left(\left|{}^1\Psi_1^2\right\rangle\right) + E\left(\left|{}^3\Psi_1^2\right\rangle\right)\right) \tag{33}$$

from which we see that the energy of the determinant $\left|\psi_1\overline{\psi}_2\right\rangle$ underestimates the energy of the pure singlet open-shell state by half the singlet-triplet energy gap.

Eq 33 suggests that we can calculate the energy of the pure singlet open-shell state using the formula:

$$E\left(\left|{}^1\Psi_1^2\right\rangle\right) = 2E\left(\left|\psi_1\overline{\psi}_2\right\rangle\right) - E\left(\left|{}^3\Psi_1^2\right\rangle\right) \tag{34}$$

In practice, since Hartree-Fock and Kohn-Sham calculations are restricted to a single determinant, we cannot calculate $E\left(\left|{}^3\Psi_1^2\right\rangle\right)$ directly, instead we can approximate the energy of the multi-determinant triplet with the energy of the single-determinant triplet :

$$E\left(\left|{}^1\Psi_1^2\right\rangle\right) \approx 2E\left(\left|\psi_1\overline{\psi}_2\right\rangle\right) - E\left(\left|\psi_1\psi_2\right\rangle\right) \tag{35}$$

This approach is usually denoted as 'sum rule'. It is rigorously correct only when all the states are written with the same set of orbitals. However, in practice it is only approximate because separate variational optimizations for the spin-mixed and triplet determinants will lead to different set of relaxed orbitals.