

RAPID COMMUNICATION | DECEMBER 16 2015

Communication: An efficient approach to compute state-specific nuclear gradients for a generic state-averaged multi-configuration self consistent field wavefunction **FREE**

Alexander A. Granovsky



J. Chem. Phys. 143, 231101 (2015)

<https://doi.org/10.1063/1.4938169>



CrossMark

Articles You May Be Interested In

Analytic energy second derivatives for general MCSCF wave functions

J. Chem. Phys. (March 1984)

A combined theoretical and experimental study of the ionic states of iodopentafluorobenzene

J. Chem. Phys. (February 2017)

Quantum firefly algorithm with stochastic search strategies

J. Appl. Phys. (August 2022)

500 kHz or 8.5 GHz?
And all the ranges in between.

Lock-in Amplifiers for your periodic signal measurements



Find out more



Communication: An efficient approach to compute state-specific nuclear gradients for a generic state-averaged multi-configuration self consistent field wavefunction

Alexander A. Granovsky^{a)}

Firefly project, Moscow, 117593 Moscow, Russian Federation

(Received 12 November 2015; accepted 7 December 2015; published online 16 December 2015)

We present a new, very efficient semi-numerical approach for the computation of state-specific nuclear gradients of a generic state-averaged multi-configuration self consistent field wavefunction. Our approach eliminates the costly coupled-perturbed multi-configuration Hartree-Fock step as well as the associated integral transformation stage. The details of the implementation within the Firefly quantum chemistry package are discussed and several sample applications are given. The new approach is routinely applicable to geometry optimization of molecular systems with 1000+ basis functions using a standalone multi-core workstation. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4938169>]

Multi-configuration self consistent field (MCSCF)^{1,2} is an important tool of modern quantum chemistry. While not accounting for dynamic electron correlation, it is one of the simplest methods capable of capturing non-dynamic correlation effects which can be very substantial in the correct description of molecular systems and especially of their excited states.^{3,4} A state-specific (SS-MCSCF) variant of MCSCF which minimizes the energy of a single configuration interaction (CI) state is often limited to modeling of a ground state. Exploration of excited states usually requires a state-averaged (SA-MCSCF) variant of MCSCF in which the minimized energy functional consists of weighted sum of energies of several CI roots. Optimization of geometry and search for conical intersections (ConInts)^{5,6} using SA-MCSCF require efficient algorithms for computation of state-specific gradients for SA-MCSCF wavefunctions.

Using the basis of arbitrary fixed spin-orbitals, electronic energy of wavefunction-based methods such as SS-MCSCF can be written in the following form:⁷

$$E = E(c, \mathbf{C}, I). \quad (1)$$

Here, c is the matrix of MO coefficients, \mathbf{C} is the vector of CI or CI-like expansion coefficients, and I denotes collectively various one and two electron integrals in the AO basis. The derivative of energy over external parameter x E^x is given by

$$E^x = \frac{\partial E}{\partial c} \cdot \frac{\partial c}{\partial x} + \frac{\partial E}{\partial \mathbf{C}} \cdot \frac{\partial \mathbf{C}}{\partial x} + \frac{\partial E}{\partial I} \cdot \frac{\partial I}{\partial x}. \quad (2)$$

For converged SS-MCSCF $\frac{\partial E}{\partial c} \equiv 0$ and $\frac{\partial E}{\partial \mathbf{C}} \equiv 0$ as c and \mathbf{C} are variationally optimized parameters of a method. Hence, the derivative expression becomes

$$E^x = \frac{\partial E}{\partial I} \cdot \frac{\partial I}{\partial x}. \quad (3)$$

This greatly simplifies the working expression for nuclear gradients of SS-MCSCF.

For the i -th state of SA-MCSCF with energy E_i and CI expansion vector \mathbf{C}_i the derivative E_i^x is

$$E_i^x = \frac{\partial E_i}{\partial c} \cdot \frac{\partial c}{\partial x} + \frac{\partial E_i}{\partial \mathbf{C}_i} \cdot \frac{\partial \mathbf{C}_i}{\partial x} + \frac{\partial E_i}{\partial I} \cdot \frac{\partial I}{\partial x}. \quad (4)$$

There are no terms $\frac{\partial E_i}{\partial c_j} \cdot \frac{\partial c_j}{\partial x}$ in (4) as $\frac{\partial E_i}{\partial c_j} \equiv 0, j \neq i$. In addition, for converged SA-MCSCF $\frac{\partial E_i}{\partial \mathbf{C}_i} = 0$ while in general $\frac{\partial E_i}{\partial c} \neq 0$.

The derivative expression is then

$$E_i^x = \frac{\partial E_i}{\partial c} \cdot \frac{\partial c}{\partial x} + \frac{\partial E_i}{\partial I} \cdot \frac{\partial I}{\partial x}. \quad (5)$$

Eq. (5) results in much more involved working expressions for nuclear gradients of SA-MCSCF as compared with those of SS-MCSCF. Indeed, due to the orbital response term we need to solve coupled-perturbed multi-configuration Hartree-Fock (CPMCHF) equations^{8,9} for z -vector.¹⁰ This requires a much larger class of transformed two-electron integrals than is needed for MCSCF itself. We note that the most efficient formulations of the MCSCF scheme like approximate second-order MCSCF by Chaban and co-authors¹¹ require per each iteration only transformed integrals over four active orbitals and intermediates obtained as a result of the first integral half-transformation, as well as a couple of Fock matrix builds. These computations can be performed very efficiently (Ref. 12). At the same time, the CPMCHF step would require transformed integrals with two virtual orbitals and two MCSCF's core orbital indices. The corresponding integral transformation is much more expensive than that needed for the MCSCF step. Depending on the details of implementation, it may require a disk-based integral sorting step. For large systems, CPMCHF integral transformation would become more time consuming than the entire MCSCF stage.

Another unpleasant feature of the CPMCHF step is the necessity to setup and solve a large system of linear equations with the detailed structure depending on the type of the

^{a)}Electronic mail: alex.granovsky@gmail.com

MCSCF problem and the basis used for CI expansion (i.e., configuration state functions vs. determinants).

Finally, it should be noted that for constant weights $\mathbf{w} = \{w_1, \dots, w_N\}$, the derivative of SA-MCSCF energy has the same functional form as Eq. (3). Indeed, we have

$$E^x = \frac{\partial E}{\partial \mathbf{w}} \cdot \frac{\partial \mathbf{w}}{\partial x} + \frac{\partial E}{\partial c} \cdot \frac{\partial c}{\partial x} + \frac{\partial E}{\partial \bar{\mathbf{C}}} \cdot \frac{\partial \bar{\mathbf{C}}}{\partial x} + \frac{\partial E}{\partial I} \cdot \frac{\partial I}{\partial x}, \quad (6)$$

where $\bar{\mathbf{C}} = \{\mathbf{C}_1, \dots, \mathbf{C}_N\}$ is a vector of CI expansion vectors. Eq. (6) reduces to Eq. (3) as $\frac{\partial E}{\partial c} = 0$, $\frac{\partial E}{\partial \bar{\mathbf{C}}} = 0$, and $\frac{\partial \mathbf{w}}{\partial x} = 0$. This means that the gradient of state-averaged energy can be obtained using the same code as that used for the state-specific MCSCF gradients. The only difference is that the state-averaged one- and two-particle density matrices as well as Lagrangian should be used instead of state-specific quantities.

Let us consider the energy functional $E(\mathbf{v}; \mathbf{x})$ where \mathbf{v} is the vector of variational parameters of a method $\mathbf{v} = \{v_1, \dots, v_M\}$ and $\mathbf{x} = \{x_1, \dots, x_K\}$ are non-variational external parameters. For SA-MCSCF, \mathbf{v} contains a matrix of MO coefficients c and CI expansion vectors $\bar{\mathbf{C}} = \{\mathbf{C}_1, \dots, \mathbf{C}_N\}$, while \mathbf{x} contains states, weights $\mathbf{w} = \{w_1, \dots, w_N\}$, molecular integrals I or, alternatively, parameters of molecular geometry and a basis set, and may contain additional parameters such as external fields and strengths. Here and throughout the communication we assume constant weights \mathbf{w} .

For SA-MCSCF with weights $\{w_1, \dots, w_N\}$, we can divide \mathbf{x} into two sets $\mathbf{x} = \{\mathbf{w}, \mathbf{y}\}$ where $\mathbf{w} = \{w_1, \dots, w_N\}$ and \mathbf{y} contains all the other non-variational parameters included into \mathbf{x} except for weights. The energy functional is then

$$E(\mathbf{v}; \mathbf{x}) = E(\mathbf{v}; \mathbf{w}, \mathbf{y}) = \sum_{i=1}^N w_i E_i(\mathbf{v}; \mathbf{y}), \quad (7)$$

where $E_i(\mathbf{v}; \mathbf{y})$ is the energy functional for the i -th state.

Let us optimize $E(\mathbf{v}; \mathbf{w}, \mathbf{y})$ while holding $\mathbf{x} = \{\mathbf{w}, \mathbf{y}\}$ intact. The values of parameters \mathbf{v} are then defined by

$$\frac{\partial E(\mathbf{v}; \mathbf{x})}{\partial \mathbf{v}} = 0, \text{ i.e., } \forall k \div 1, M \frac{\partial E(\mathbf{v}; \mathbf{x})}{\partial v_k} = 0. \quad (8)$$

Eq. (8) defines optimal \mathbf{v} as a function of $\tilde{\mathbf{v}}(\mathbf{x})$ such that

$$\left. \frac{\partial E(\mathbf{v}; \mathbf{x})}{\partial \mathbf{v}} \right|_{\mathbf{v}=\tilde{\mathbf{v}}(\mathbf{x})} = \frac{\partial E(\tilde{\mathbf{v}}(\mathbf{x}); \mathbf{x})}{\partial \mathbf{v}} = 0. \quad (9)$$

Optimal $\tilde{\mathbf{v}}$ defines function $\tilde{E}(\mathbf{x})$ as follows:

$$\tilde{E}(\mathbf{x}) = E(\tilde{\mathbf{v}}(\mathbf{x}); \mathbf{x}) \quad (10)$$

Differentiating $\tilde{E}(\mathbf{x})$ and using Eq. (9), we get

$$\begin{aligned} \frac{\partial \tilde{E}(\mathbf{x})}{\partial \mathbf{x}} &= \frac{\partial E(\tilde{\mathbf{v}}(\mathbf{x}); \mathbf{x})}{\partial \mathbf{x}} + \left(\frac{\partial E(\tilde{\mathbf{v}}(\mathbf{x}); \mathbf{x})}{\partial \mathbf{v}}, \frac{\partial \tilde{\mathbf{v}}(\mathbf{x})}{\partial \mathbf{x}} \right) \\ &= \frac{\partial E(\tilde{\mathbf{v}}(\mathbf{x}); \mathbf{x})}{\partial \mathbf{x}}, \end{aligned} \quad (11)$$

i.e., the analogue of Hellmann-Feynman theorem. Using Eq. (11), for energy functional of Eq. (7) we get

$$\frac{\partial \tilde{E}(\mathbf{w}, \mathbf{y})}{\partial w_k} = \frac{\partial E(\tilde{\mathbf{v}}(\mathbf{w}, \mathbf{y}); \mathbf{w}, \mathbf{y})}{\partial w_k} = E_k(\tilde{\mathbf{v}}(\mathbf{w}, \mathbf{y}); \mathbf{y}). \quad (12)$$

Differentiating Eq. (12) and using commutativity of partial derivatives, we obtain

$$\begin{aligned} \frac{\partial E_k(\tilde{\mathbf{v}}(\mathbf{w}, \mathbf{y}); \mathbf{y})}{\partial y_j} &= \frac{\partial^2 E(\tilde{\mathbf{v}}(\mathbf{w}, \mathbf{y}); \mathbf{w}, \mathbf{y})}{\partial y_j \partial w_k} \\ &= \frac{\partial}{\partial w_k} \left(\frac{\partial E(\tilde{\mathbf{v}}(\mathbf{w}, \mathbf{y}); \mathbf{w}, \mathbf{y})}{\partial y_j} \right). \end{aligned} \quad (13)$$

Assuming y_j in Eq. (13) to be the nuclear coordinate variable, we finally get

$$g_k = \frac{d\bar{g}}{dw_k}, \quad (14)$$

where g_k is the state-specific gradient of k -th state and \bar{g} is a gradient of the state-averaged energy. Eqs. (12)–(14) are the central results of our communication.

The standard expression for a component of the nuclear gradient of electronic energy E^x in molecular gradient theory is as follows:¹³⁻¹⁵

$$E^x = \text{Tr}(PH_1^x) + \text{Tr}(WS^x) + \sum_{\mu\nu\lambda\sigma} \Gamma^{\mu\nu\lambda\sigma} \cdot (\mu\nu|\lambda\sigma)^x. \quad (15)$$

Here, P is the response type one-particle density matrix, W is the so-called energy-weighted density matrix, and Γ is the two-particle response type density matrix. H_1^x and S^x are the derivatives of one electron Hamiltonian and the overlap matrix, and $(\mu\nu|\lambda\sigma)^x$ is the two-electron integral derivative. For fully variational theories like SS-MCSCF and for state-averaged energy of a SA-MCSCF, P and Γ become standard expectation value type density matrices.

Differentiating Eq. (15) with respect to w_i , we obtain

$$\begin{aligned} E_i^x &= \frac{dE^x}{dw_i} = \text{Tr} \left(\frac{dP}{dw_i} \cdot H_1^x \right) + \text{Tr} \left(\frac{dW}{dw_i} \cdot S^x \right) \\ &\quad + \sum_{\mu\nu\lambda\sigma} \frac{d\Gamma^{\mu\nu\lambda\sigma}}{dw_i} \cdot (\mu\nu|\lambda\sigma)^x. \end{aligned} \quad (16)$$

The formal expression for E_i^x is

$$E_i^x = \text{Tr}(P_i H_1^x) + \text{Tr}(W_i S^x) + \sum_{\mu\nu\lambda\sigma} \Gamma_i^{\mu\nu\lambda\sigma} \cdot (\mu\nu|\lambda\sigma)^x, \quad (17)$$

where P_i , W_i , and Γ_i are state-specific response-type density matrices.

Comparing Eqs. (16) and (17) and taking into account that Eqs. (14) and (15) are to be satisfied for arbitrary H_1^x , S^x , and $(\mu\nu|\lambda\sigma)^x$ we arrive at

$$P_i = \frac{dP}{dw_i}, W_i = \frac{dW}{dw_i}, \Gamma_i = \frac{d\Gamma}{dw_i}. \quad (18)$$

In other words, the weight derivatives of state-averaged SA-MCSCF density matrices are equal to response-type state-specific SA-MCSCF density matrices.

To conclude, we note that in virtue of Eq. (18) any state-specific one and two electron property for SA-MCSCF wavefunctions can be obtained as a corresponding weight derivative.

Considered formally, Eq. (14) is at least as complicated as Eq. (5) (Ref. 16). The key idea behind its practical use is that the required derivatives can be computed numerically. This is a general approach that can be used not only for gradients but also for Hessians, electrical moments, polarizabilities, etc., as well as directly for density matrices.

This approach was implemented in Firefly Quantum Chemistry package (Ref. 17) back in 2008 (Ref. 18). By default, our implementation uses three-point second order finite differences to compute state-specific gradients. Firefly assumes that SA-MCSCF weights sum up to unity. To avoid problems imposed by this assumption, in practical applications we perturb the weights, back-renormalize them, perform computations, and finally renormalize the answer. Tests showed that a finite difference step Δw of 0.001–0.002 is usually enough to compute gradients with a precision of 5–6 decimal digits. This precision was found to be sufficient for geometry optimization, location of ConInts, and even for semi-numerical Hessians.

As compared to the traditional approach, our scheme requires three SA-MCSCF energy computations and two averaged energy gradient computations when using central differences (as a gradient at the central point does not enter the finite difference equations). In the case of non-symmetric finite differences, an additional energy gradient computation is needed. Two of the three SA-MCSCF computations start with almost converged orbitals from the first point thus requiring only a few iterations to converge. We need to increase the overall precision of computations by an order or two to obtain reliable finite differences. It should be noted that the increased precision would be required for the standard approach as well, as SA-MCSCF state energies are not fully variational.

The central finite difference scheme is not applicable in the vicinity of ConInts because it is no longer numerically stable. Indeed, in the vicinity of state crossings SA-MCSCF may have qualitatively different solutions for slightly differently “weighted” setups. Let us consider a two-state quasi-degenerate system with equal unperturbed weights as an example. If one increases the weight of the upper state, it would be described somewhat better than the lower state. As a result, two states may flip. We thus conclude that for numerical stability we should only increase the weight of the lower state as demonstrated below:

step 1: $w_1 = 0.5, w_2 = 0.5$,

step 2: $w_1 = 0.5 + \Delta w, w_2 = 0.5$,

step 3: $w_1 = 0.5 + 2 \cdot \Delta w, w_2 = 0.5$.

In addition, Firefly uses state-tracking when needed and is capable of reassigning states on the fly.

As for conical intersections, we emphasize the important special case of two-state state-averaged MCSCF (SA2-MCSCF). When searching for ConInts, gradients of both states are needed simultaneously. This does not however impose any additional costs because upon differentiating Eq. (7) we arrive at

$$\sum_i w_i g_i = \bar{g} \quad (19)$$

and thus both gradients can be computed from a single set of computations as \bar{g} is known.

The advantage of the suggested method is that it is equally applicable to standard MCSCF and to density fitting (DF) MCSCF implementations.¹⁹ Moreover, it does not require any information on the details of the underlying MCSCF being applicable to complete active space self-consistent field

(CASSCF),^{20,21} restricted active space self consistent field (RASSCF),²² generic occupation restricted multiple active spaces (ORMAS),^{23,24} or any other type MCSCF.

The CPMCHF step and the associated integral transformation stage are completely eliminated in our approach. While CPMCHF was found to be of minor costs for DF-SA-MCSCF gradients,^{25,26} it is still the most time consuming part of conventional SA-MCSCF gradient computations.²⁵ This makes our approach the fastest canonical SA-MCSCF gradient method for virtually any molecular system and also very modest in its memory requirements, and we are presently working on the extension of our approach for the computation of non-adiabatic coupling vectors.^{5,6} As sample applications we cite the work of Bochenkova *et al.* who computed ConInts in a molecule of 4-hydroxybenzylidene-1,2-dimethylimidazoline (HBDI) - the model chromophore of a green fluorescent protein (GFP). Calculations involved 315 basis functions at CASSCF(14e,13o) level of theory (Ref. 27). Then, CASSCF's ConInts were partially refined using XMCQDPT2 (Ref. 28) energies. We also mention the work of Ioffe *et al.* who computed equilibrium geometries of three lowest excited states of a fulvene-like molecule, 1-butyl-4-(1H-inden-1-ylidene)-1,4-dihydropyridine (BIDP) (646 basis functions, CASSCF(12e,12o)). The S_0/S_1 ConInt structure was located as well (Ref. 29). Finally, we cite the work of Domratcheva *et al.* who have successfully performed SA-MCSCF optimizations with more than 1300 basis functions using our code. The active site of the bacterial blue-light sensing protein BLUF (stands for the sensor of Blue Light Using Flavin adenine dinucleotide) containing the flavin chromophore and several amino acid residues surrounding it was considered. Geometry optimization of the rather large active site model in the ground state at the CASSCF(6e,4o) level of theory with state averaging over the four states was carried out (Ref. 30). At optimized geometry, the excitation spectrum was then computed at the XMCQDPT2 level of theory.

ACKNOWLEDGMENTS

The author is very grateful to Dr. Vladimir I. Pupyshev for multiple stimulating discussions. This work was partially supported by the RFBR (Project No. 14-03-00887).

¹J. Frenkel, *Wave Mechanics, Advanced General Theory* (Oxford University Press, Oxford, 1934), pp. 460–462.

²D. R. Hartree, W. Hartree, and B. Swirles, *Philos. Trans. R. Soc., A* **238**, 229 (1939).

³M. W. Schmidt and M. S. Gordon, *Annu. Rev. Phys. Chem.* **49**, 233 (1998).

⁴V. Veryazov, P. Å. Malmqvist, and B. O. Roos, *Int. J. Quantum Chem.* **111**, 3329 (2011).

⁵E. Teller, *J. Phys. Chem.* **41**, 109 (1937).

⁶G. Herzberg and H. C. Longuet-Higgins, *Discuss. Faraday Soc.* **35**, 77 (1963).

⁷E. A. Salter, G. W. Trucks, and R. J. Bartlett, *J. Chem. Phys.* **90**, 1752 (1989).

⁸Y. Osamura, Y. Yamaguchi, and H. F. Schaefer, *J. Chem. Phys.* **77**, 383 (1982).

⁹M. R. Hoffmann, D. J. Fox, J. F. Gaw, Y. Osamura, Y. Yamaguchi, R. S. Grev, G. Fitzgerald, H. F. Schaefer, P. J. Knowles, and N. C. Handy, *J. Chem. Phys.* **80**, 2660 (1984).

¹⁰N. C. Handy and H. F. Schaefer III, *J. Chem. Phys.* **81**, 5031 (1984).

¹¹G. Chaban, M. W. Schmidt, and M. S. Gordon, *Theor. Chem. Acc.* **97**, 88 (1997).

- ¹²A. A. Granovsky, (2005), <http://classic.chem.msu.su/gran/gamess/gcmcsf.pdf>, accessed October 2015.
- ¹³J. E. Rice and R. D. Amos, *Chem. Phys. Lett.* **122**, 585 (1985).
- ¹⁴E. A. Salter, G. W. Trucks, G. Fitzgerald, and R. J. Bartlett, *Chem. Phys. Lett.* **141**, 61 (1987).
- ¹⁵G. W. Trucks, E. A. Salter, C. Sosa, and R. J. Bartlett, *Chem. Phys. Lett.* **147**, 359 (1988).
- ¹⁶J. Stålring, A. Bernhardsson, and R. Lindh, *Mol. Phys.* **99**, 103 (2001).
- ¹⁷A. A. Granovsky, Firefly version 8, Firefly Project, 2013, <http://classic.chem.msu.su/gran/firefly/index.html>.
- ¹⁸A. A. Granovsky, (2008), <http://classic.chem.msu.su/gran/gamess/ss-gradients.pdf>, accessed October 2015.
- ¹⁹F. Aquilante, T. B. Pedersen, R. Lindh, B. O. Roos, A. S. de Merás, and H. Koch, *J. Chem. Phys.* **129**, 024113 (2008).
- ²⁰L. M. Cheung, K. R. Sundberg, and K. Ruedenberg, *J. Am. Chem. Soc.* **100**, 8024 (1978).
- ²¹B. O. Roos, *Adv. Chem. Phys.* **69**, 399 (1987).
- ²²J. Olsen, B. O. Roos, P. Jørgensen, and H. J. Aa. Jensen, *J. Chem. Phys.* **89**, 2185 (1988).
- ²³J. Ivanic, *J. Chem. Phys.* **119**, 9364 (2003).
- ²⁴J. Ivanic, *J. Chem. Phys.* **119**, 9377 (2003).
- ²⁵M. G. Delcey, T. B. Pedersen, F. Aquilante, and R. Lindh, *J. Chem. Phys.* **143**, 044110 (2015).
- ²⁶J. W. Snyder, Jr., E. G. Hohenstein, N. Luehr, and T. J. Martínez, *J. Chem. Phys.* **143**, 154107 (2015).
- ²⁷A. V. Bochenkova and L. H. Andersen, *Faraday Discuss.* **163**, 297 (2013).
- ²⁸A. A. Granovsky, *J. Chem. Phys.* **134**, 214113 (2011).
- ²⁹I. Ioffe, A. L. Dobryakov, A. A. Granovsky, N. P. Ernsting, and J. L. P. Lustres, *ChemPhysChem* **12**, 1860 (2011).
- ³⁰A. Udvarhelyi, M. Olivucci, and T. Domratcheva, *J. Chem. Theory Comput.* **11**, 3878 (2015).