Quantum Monte Carlo for electronic structure: excited-state calculations using linear response theory

Bastien Mussard[†], **Emanuele Coccia**[*,**,‡], Roland Assaraf[**], Matt Otten[***], Cyrus Umrigar[***] and Julien Toulouse[**]

[†] Laboratoire de Chimie Théorique & Institut du Calcul et de la Simulation, Université Pierre et Marie Curie, Paris (France)
[*] Centro S3 - Istituto di Nanoscienze, CNR, Modena, Italy
[**] Laboratoire de Chimie Théorique, Université Pierre et Marie Curie & CNRS, Paris (France)
[‡] Dipartimento di Scienze Fisiche e Chimiche, Universitá degli Studi dell'Aquila, L'Aquila (Italy)
[***] Physics Department, LASSP, Cornell University, Ithaca (USA)

RPA Workshop 2017

B. Mussard, **EC**, R. Assaraf, M. Otten, C. Umrigar and J. Toulouse, Adv. Quantum Chem., submitted

Outline

1. Quantum Monte Carlo: overview
2. Wave function optimization: linear method
3. Linear-response theory: QMC excited states
4. Test case: Be atom
5. Conclusions

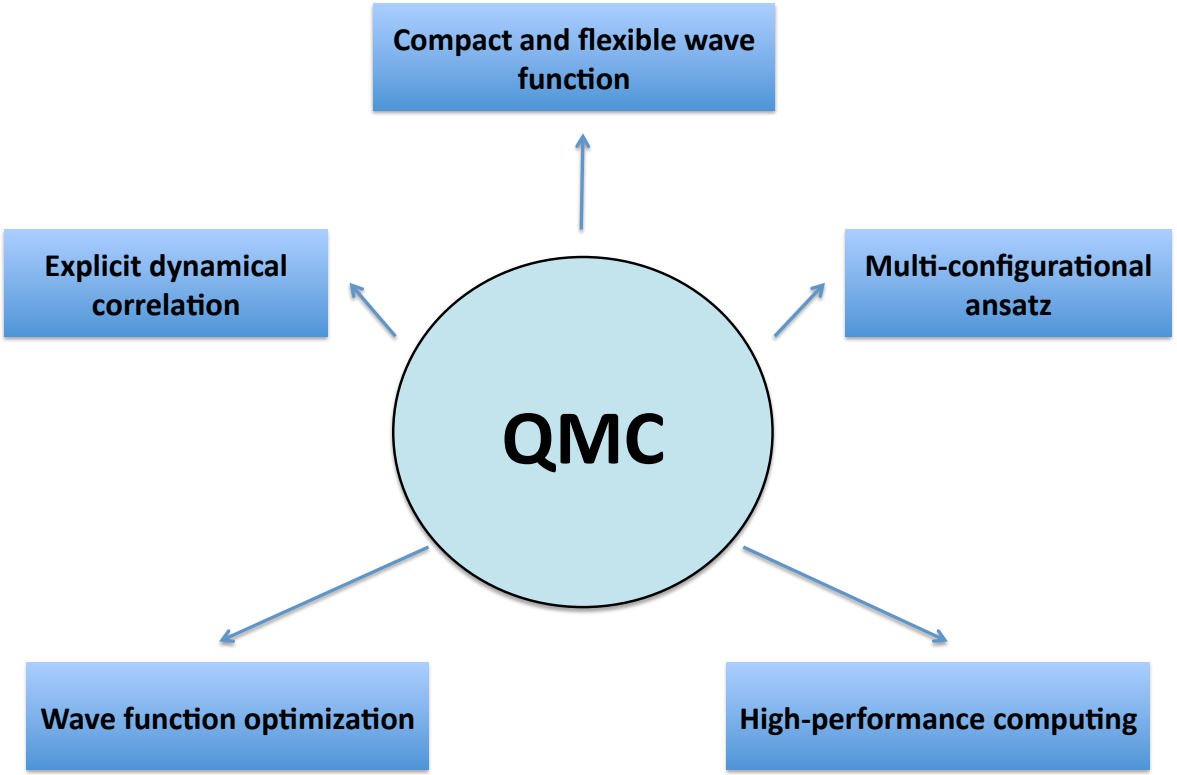
Methods for electronic structure

From the application standpoint:

- Wave function methods (Hartree-Fock, Perturbation theory, Configuration Interaction, Coupled Cluster): high accuracy, (usually) large computational cost
- Density Functional Theory (DFT): tradeoff between accuracy and cost, good scalability (N^3)
- "Third way": Quantum Monte Carlo (QMC) methods, embarrassing parallelism

QMC suitable for massive calculations in chemistry (hundreds of electrons)

Why Quantum Monte Carlo



- Alternative to DFT and quantum chemistry for ground-state properties
- Cons: large prefactor, error, no "black-box"
- Extension of QMC to excited states?

QMC for excited states: state of the art

- For lowest energy state of a given symmetry: same as ground state calculation
- For a state that is not the lowest one in a given symmetry:
  - state-average approach: minimization of a weighted average of the energies of S states (F. Schautz and C. Filippi, *J. Chem. Phys.*, **120**, 10931 (2004))
  - state-specific approach: minimization of the energy of the targeted state (P. M. Zimmerman *et al.*, *J. Chem. Phys.*, **131**, 124103 (2009))
- Variational approach for the Antisymmetrized Geminal Power ansatz (E. Neuscamman, *J. Chem. Phys.*, **145**, 081103 (2016), L. Zhao and E. Neuscamman, *J. Chem. Theory Comput.* **12**, 3719 (2016), L. Zhao and E. Neuscamman, *J. Chem. Theory Comput.*, **12**, 3436 (2016))

Variational Monte Carlo

$$E_{VMC} = \min_{\mathbf{p}} \frac{\langle \Psi(\mathbf{p}) | \hat{H} | \Psi(\mathbf{p}) \rangle}{\langle \Psi(\mathbf{p}) | \Psi(\mathbf{p}) \rangle} \geq E_{\text{exact}}$$

- Integration in the $3N$ variational space
- With M sampling points the error $\epsilon \propto \frac{1}{\sqrt{M}}$
- M points randomly drawn from $\Pi \equiv \frac{|\Psi|^2}{\int d\mathbf{R} |\Psi|^2}$
- Choice of the trial wave function Ψ

Trial wave function

Jastrow-Slater wave function

$$|\Psi(\mathbf{p})\rangle = \hat{J}(\boldsymbol{\alpha}) \sum_{I=1}^{N_{\text{CSF}}} c_I \, |C_I\rangle$$

- $\hat{J}(\boldsymbol{\alpha})$: Jastrow factor (with e-e, e-n, e-e-n terms)
- $|C_I\rangle$: configuration state function (CSF)
- Slater determinants are made of orbitals ϕ expanded on a Slater basis:

$$\phi_k(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{basis}}} \lambda_{k\mu} \chi_\mu(\mathbf{r})$$

$$\chi(\mathbf{r}) = N(\zeta)\, r^{n-1}\, e^{-\zeta r}\, S_{l,m}(\theta, \phi)$$

- Parameters to optimize $\mathbf{p} = \{\boldsymbol{\alpha}, \mathbf{c}, \boldsymbol{\lambda}, \boldsymbol{\zeta}\}$

J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **126** 084102 (2007)

J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **128** 174101 (2008)

Wave function optimization

- QMC affected by:
  - the systematic error
  - the statistical uncertainty
- QMC wave function optimization:
  - Noise affecting the minimum search
  - Naive energy minimization on a finite Monte Carlo sample
  - Variance minimization: bound from below (~ 100 parameters)
  - Efficient energy minimization: optimization of ~ 1000 parameters (including non-linear ones) independently of the Monte Carlo sample

J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **126** 084102 (2007)
C. Umrigar *et al.*, *Phys. Rev. Lett.*, **98** 110201 (2007)
J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **128** 174101 (2008)

Linear method

- Linear order expansion of Ψ around $\mathbf{p}^0$ in $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^0$:

$$|\Psi_{\text{lin}}(\mathbf{p})\rangle = |\Psi_0\rangle + \sum_i \Delta p_i \, |\Psi_i\rangle$$

- $|\Psi_0\rangle = |\Psi(\mathbf{p}^0)\rangle$ and $|\Psi_i\rangle = \dfrac{\partial \, |\Psi(\mathbf{p}^0)\rangle}{\partial p_i}$
- Normalization of wave function chosen so that the derivatives $|\Psi_i\rangle$ are orthogonal to $|\Psi_0\rangle$
- Minimization of the energy $\rightarrow$ generalized eigenvalue equation:

$$\begin{pmatrix} E_0 & \mathbf{g}^{\mathsf{T}}/2 \\ \mathbf{g}/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^{\mathsf{T}} \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix}$$

- $E_0 = \langle\Psi_0| \, \hat{H} \, |\Psi_0\rangle$, $g_i = \langle\Psi_i| \, \hat{H} \, |\Psi_0\rangle$, $H_{ij} = \langle\Psi_i| \, \hat{H} \, |\Psi_j\rangle$, $S_{ij} = \langle\Psi_i|\Psi_j\rangle$
- **Update of the parameters: $\mathbf{p}^0 \rightarrow \mathbf{p}^0 + \Delta\mathbf{p}$**

C. Umrigar *et al.*, *Phys. Rev. Lett.*, **98** 110201 (2007), J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **128** 174101 (2008)

Linear method on a finite VMC sample

The generalized eigenvalue equation is estimated as

$$\begin{pmatrix} E_0 & \mathbf{g}_R^{\mathsf{T}}/2 \\ \mathbf{g}_L/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^{\mathsf{T}} \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix}$$

with

$$g_{L,i}/2 = \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{H(\mathbf{R})\Psi_0(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2} \qquad g_{R,j}/2 = \left\langle \frac{\Psi_0(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{H(\mathbf{R})\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2}$$

$$H_{ij} = \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{H(\mathbf{R})\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2} \qquad S_{ij} = \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2}$$

non-symmetric!

$\Longrightarrow$ Zero-variance principle (PRL, **87** 043401 (2001):

If there are some $\Delta\mathbf{p}$ so that $\Psi_0(\mathbf{R}) + \sum_j \Delta p_j\, \Psi_j(\mathbf{R}) = \Psi_{\text{exact}}(\mathbf{R})$
then $\Delta\mathbf{p}$ are found with **zero variance**

Non-symmetric estimators reduce the fluctuations on $\Delta\mathbf{p}$ by 1 or 2 orders of magnitude

C. Umrigar *et al.*, *Phys. Rev. Lett.*, **98** 110201 (2007)

Linear-response theory

- Excitation energies from linear-response theory
- Time-dependent perturbation

$$\hat{H}(t) = \hat{H} + \gamma \hat{V}(t)$$

- Time evolution of $\Psi(\mathbf{p}(t))$ according to the Dirac-Frenkel variational principle:

$$\frac{\partial}{\partial p_i^*} \langle \Psi(\mathbf{p}(t)) | \hat{H}(t) - i \frac{\partial}{\partial t} | \Psi(\mathbf{p}(t)) \rangle / \langle \Psi(\mathbf{p}(t)) | \Psi(\mathbf{p}(t)) \rangle = 0$$

- Wave function expansion around $\mathbf{p}^0 = \mathbf{p}(t_0)$

$$\begin{aligned} \Psi(\mathbf{p}(t)) &= \Psi(\mathbf{p}^0) + \sum_i \Delta p_i(t) \frac{\partial \Psi(\mathbf{p}^0)}{\partial p_i} \\ &+ \frac{1}{2} \sum_{ij} \Delta p_i(t) \Delta p_j(t) \frac{\partial^2 \Psi(\mathbf{p}^0)}{\partial p_i \partial p_j} + \ldots \end{aligned}$$

- $\Delta \mathbf{p}(t) = \mathbf{p}(t) - \mathbf{p}^0$

R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press (1992)

Linear-response theory

- Infinitesimal perturbation $\gamma \to 0$
- First-order variations $\Delta\mathbf{p}$ retained

$$\mathbf{A} \cdot \Delta\mathbf{p}(t) + \mathbf{B} \cdot \Delta\mathbf{p}(t)^* = i\mathbf{S}\frac{\partial \Delta\mathbf{p}(t)}{\partial t}$$

- with $A_{ij} = \langle \frac{\partial \Psi}{\partial p_i} | \hat{H} - E_0 | \frac{\partial \Psi}{\partial p_j} \rangle$, $B_{ij} = \langle \frac{\partial^2 \Psi}{\partial p_i \partial p_j} | \hat{H} | \Psi \rangle$, $S_{ij} = \langle \frac{\partial \Psi}{\partial p_i} | \frac{\partial \Psi}{\partial p_j} \rangle$
- Free-oscillation solutions

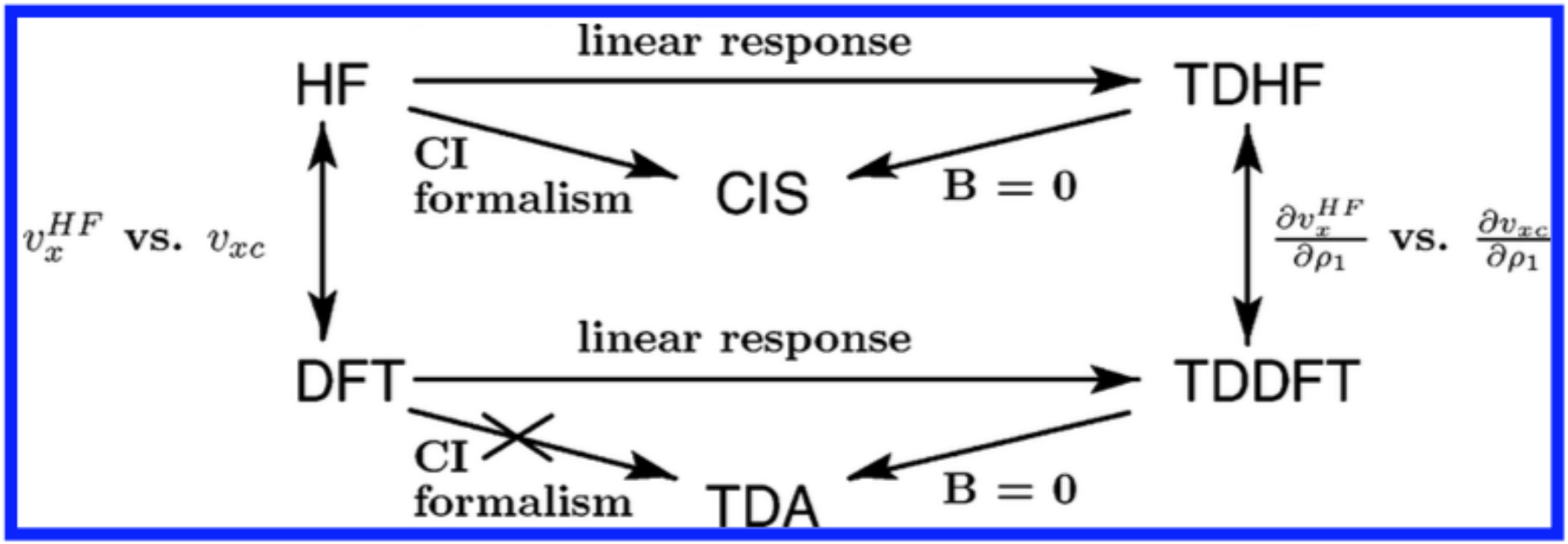
$$\Delta\mathbf{p}(t) = \mathbf{X}e^{-i\omega t} + \mathbf{Y}^* e^{i\omega t}$$

- $\omega \to$ excitation energy
- Generalized eigenvalue equation:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & -\mathbf{S}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$$

R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press (1992)

Linear-response theory



- Applied to HF or MCSCF wave functions
- Casida equations are the TDDFT version of the linear-response eigenvalue equation
- $\mathbf{B} = 0 \rightarrow$ Tamm-Dancoff approximation (TDA)

$$\mathbf{AX} = \omega \mathbf{SX}$$

A. Dreuw and M. Head-Gordon, *Chem. Rev.*, **105**, 4009 (2005)

Excited states within VMC (LR-VMC)

Linear method

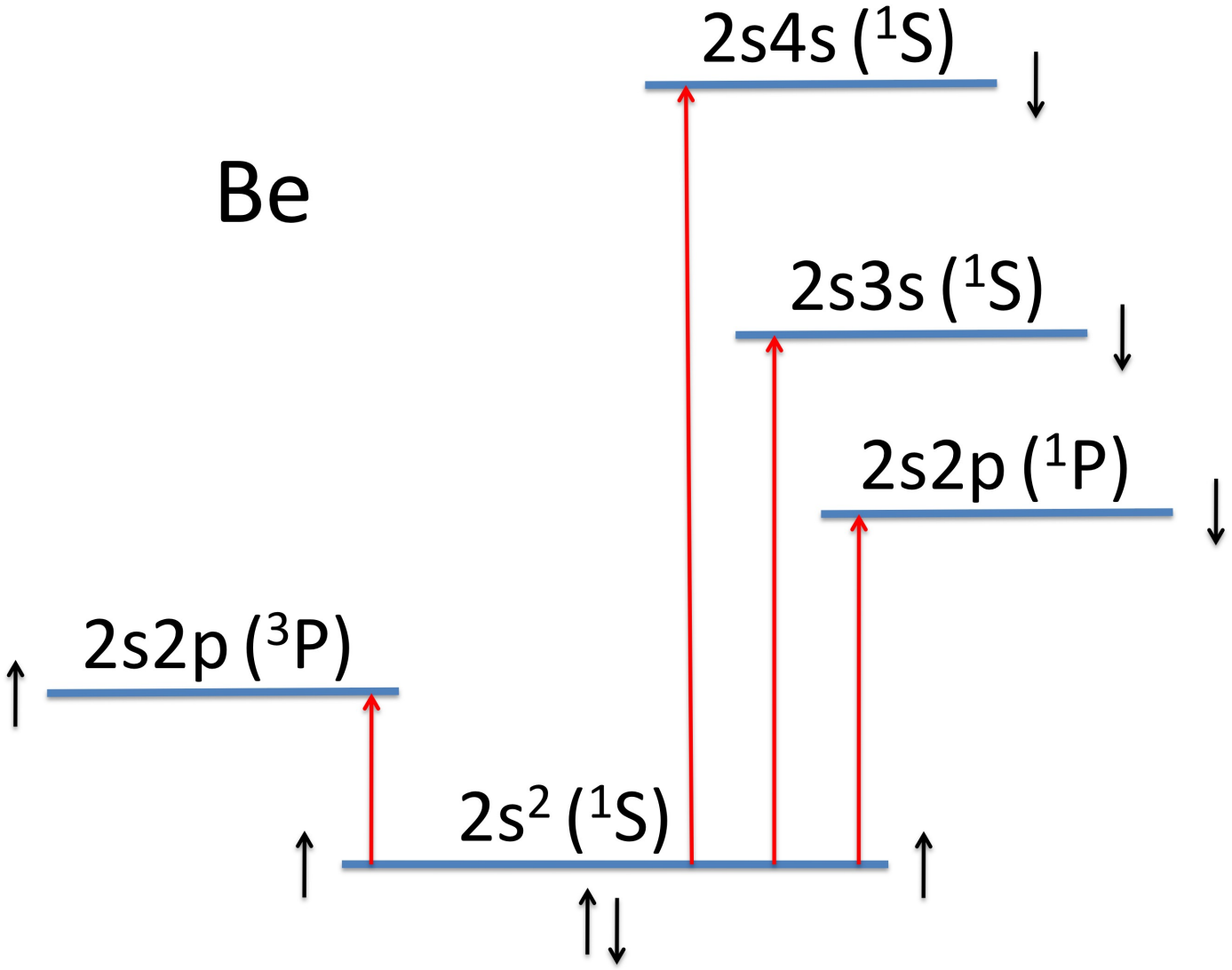
- Wf optimization
- Free choice of Ψ in VMC
- Only first derivatives
- Approximation to excitation energies
- $\mathbf{H}\Delta\mathbf{p} = (E_{\text{lin}} - E_0)\mathbf{S}\Delta\mathbf{p}$

Linear-response theory

- Excited-state approach
- Applied to HF, MCSCF, DFT
- First (TDA) and second derivatives
- ω excitation energies
- $\mathbf{AX} = \omega\mathbf{SX}$

- $\mathbf{H}$, $\mathbf{S}$ and $\Delta\mathbf{p}$ defined in the space of $\{\Psi_i\}$ (at minimum, gradient zero)
- For a HF ground-state, LR-VMC beyond CIS (thanks to the Jastrow and optimization)

Beryllium atom



Be atom: $2s3s(^1S)$ excitation

- Single-determinant wave function
- 10^5 blocks $\times$ 10^4 steps per block / energies in Hartree
- Slater basis sets (I. Ema *et al.*, J. Comput. Chem., **24**, 859 (2003)):
  - VB1 $[5s, 1p]$
  - VB2 $[6s, 2p, 1d]$
- Selective contribution from parameters:
  - j $\rightarrow$ only Jastrow
  - j+o $\rightarrow$ Jastrow and orbitals

Ha	CIS	j	j+o	Exp.
VB1	0.378	0.3147(1)	0.2713(2)	0.249
VB2	0.228	0.3194(1)	0.247(4)	0.249

- Error LR-VMC/j+o (VB2) 0.05(11) eV
- Effect of symmetrising **A**

Ha	10^2 steps	10^4 steps
Asymm	0.271(1)	0.2713(2)
Symm	0.270(6)	0.2713(2)

Exp: A. Kramida and W. C. Martin, *J. Phys. Chem. Ref. Data*, **26** 1185 (1997)

Be atom: other excitations

- $2s4s(^1S)$ excitation
- Error LR-VMC/j+o (VB2) 1.4(2) eV

Ha	CIS	j	j+o	Exp.
VB1	2.639	0.7415(2)	0.6380(5)	0.297
VB2	0.470	0.7506(2)	0.348(9)	0.297

- $2s2p(^1P)$ excitation (first singlet excited state)
- Error LR-VMC/j+o (VB2) 0.14(3) eV

Ha	CIS	j	j+o	Exp.
VB1	0.220	0.3147(1)	0.1990(6)	0.194
VB2	0.189	0.3194(1)	0.189(1)	0.194

- $2s2p(^3P)$ triplet excitation
- Error LR-VMC/j+o (VB2) 0.2(2) eV

Ha	CIS	j	j+o	Exp.
VB1	0.068	0.3147(1)	0.0925(4)	0.100
VB2	0.063	0.3194(1)	0.0925(6)	0.100

Conclusions

- VMC implementation of linear response for excited-state calculations
- Formal analogy between the linear method and linear-response theory
- Results for the beryllium atom:
  - role of the different set of parameters
  - several excitations studied ($2s2p$, $2s3s$, $2s4s$)

Ongoing work:

- Larger basis set (VB3)
- Optimizing exponents
- Convergence with respect to CSFs
- Application to other atoms and molecules
- Beyond TDA, i. e. including second derivatives
- Properties other than excitation energies

Acknowledgements

- Julien Toulouse (LCT, UPMC, France)
- Bastien Mussard (ICS, UPMC, France)
- Cyrus Umrigar (Cornell, US)
- Roland Assaraf (LCT, UPMC, France)
- Eleonora Luppi (LCT, UPMC, France)
- Leonardo Guidoni (Univaq, Italy)

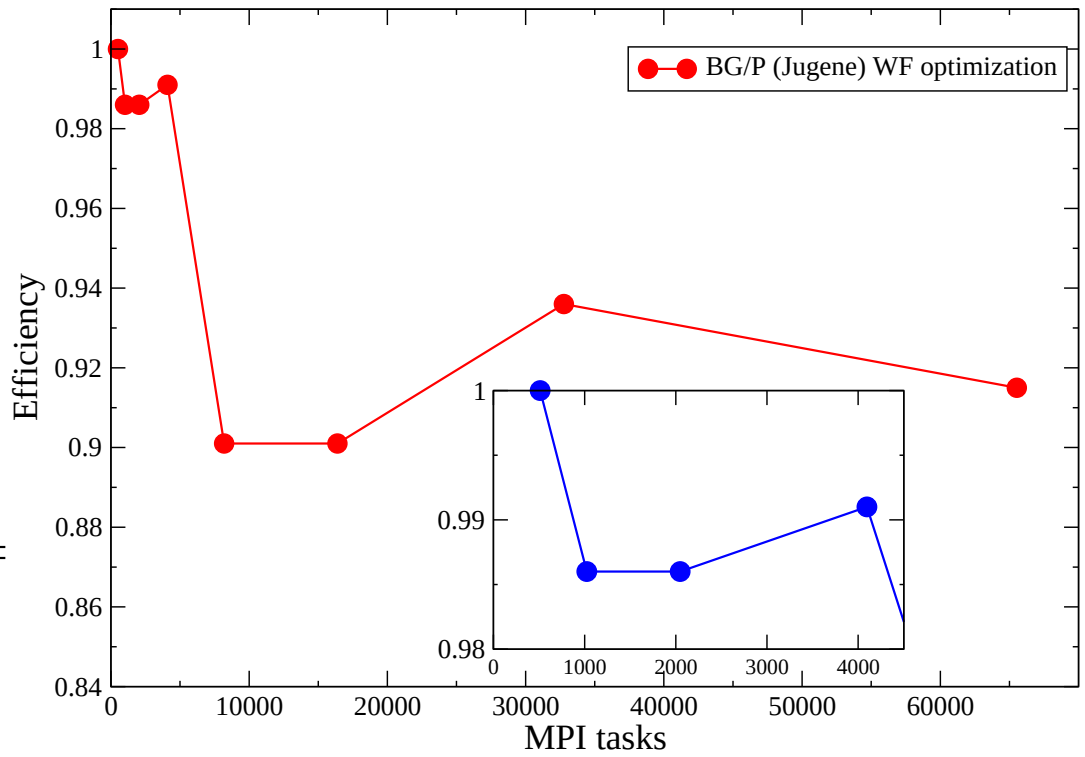
QMC for electronic ground state

- Geometry optimization
  - M. Barborini and L. Guidoni, J.Chem.Theory Comput., vol. 8, 1260 (2012)
  - **EC** and L. Guidoni, J. Comput. Chem., vol. 33, 2332 (2012)
  - **EC**, D. Varsano and L. Guidoni, J. Chem. Theory Comput., vol. 10, 501 (2014)
- Electronic properties
  - **EC**, O. Chernomor, M. Barborini, S. Sorella and L. Guidoni, J. Chem. Theory Comput, vol. 8, 1952 (2012)
  - M. Barborini and **EC**, J. Chem. Theory Comput, vol. 11, 5696 (2015)
- Vibrational properties
  - A. Zen, Y. Luo, S. Sorella and L. Guidoni, J. Chem. Theory Comput., vol. 9, 8 (2013)
- Multiscale QMC/MM
  - **EC**, D. Varsano and L. Guidoni, J. Chem. Theory Comput., vol. 9, 8 (2013)
- Chemical reactions
  - M. Barborini and L. Guidoni, J. Chem. Phys., vol. 137, 224309 (2012)
  - S. Chu, **EC**, M. Barborini and L. Guidoni, J. Chem. Theory Comput., vol. 12, 5803 (2016)

High-performance computing with QMC

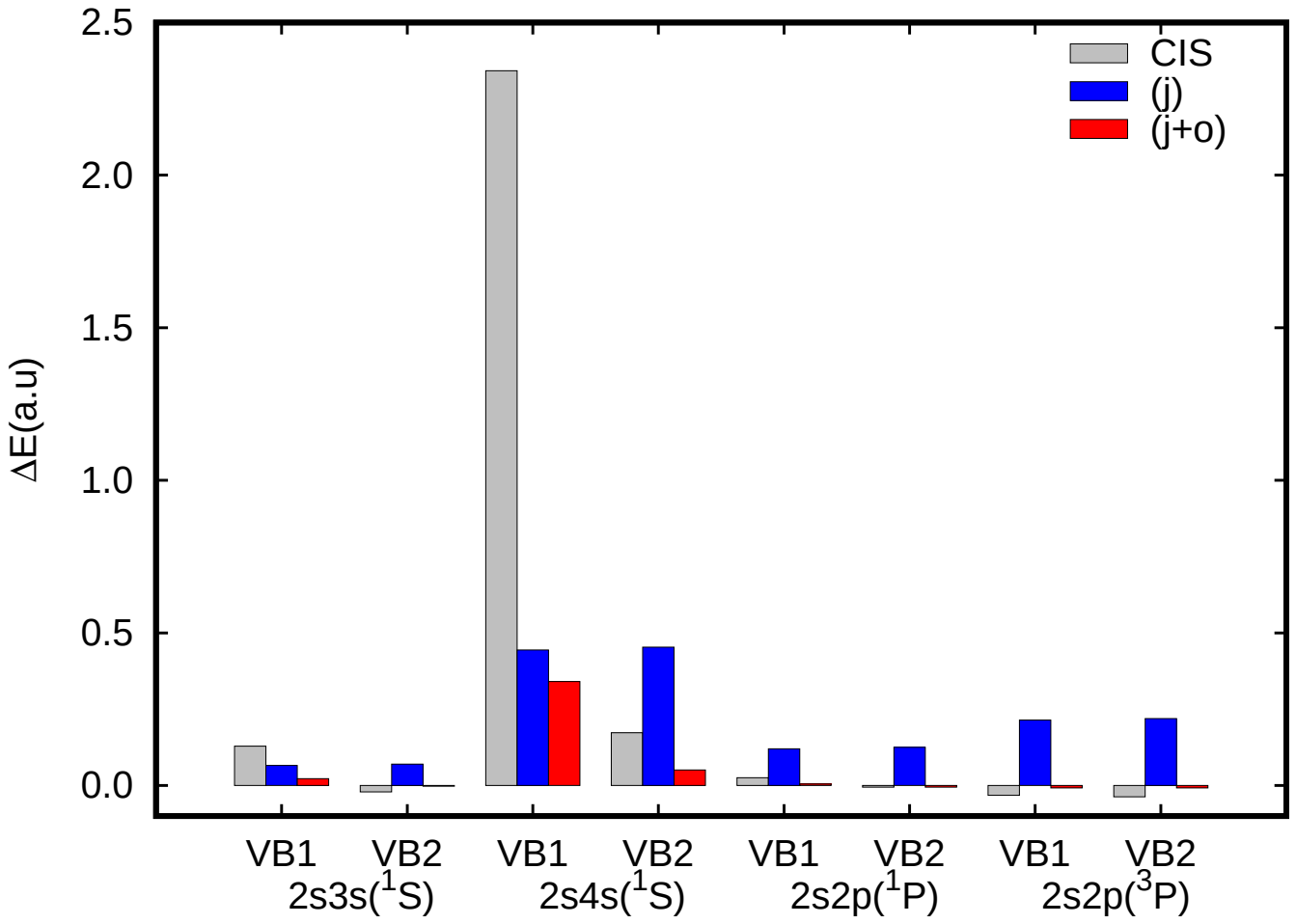
- QMC embarrassing parallel algorithm
- Wave function optimization of a biochromophore (120 valence electrons)
- Pure MPI runs
- Efficiency=T*/T(MPI tasks), where T*=T(512)

MPI tasks	time (s)	Efficiency
512	1602	-
1024	1624	0.986
2048	1624	0.986
4096	1617	0.991
8192	1779	0.901
16384	1769	0.901
32768	1712	0.936
65536	1751	0.915



EC and L. Guidoni, *J. Comput. Chem.*, **33**, 2332 (2012)

Excitations for Be



Be atom: $2s3s(^1S)$ excitation

- CAS(2,4) wave function
- 10^5 blocks $\times$ 10^4 steps per block
- Excitation energies in Hartree
- Slater basis sets:
  - VB1 $[5s, 1p]$
  - VB2 $[6s, 2p, 1d]$
- Selective contribution from parameters:
  - j $\rightarrow$ only Jastrow
  - j+c+o $\rightarrow$ Jastrow, CSF and orbitals
  - j+o $\rightarrow$ Jastrow and orbitals

Ha	MCSCF	j	j+c	j+c+o	Exp.
VB1	0.378	0.3156(2)	0.3235(2)	0.3255(1)	0.249
VB2	0.228	0.3244(1)	0.3226(2)	0.250(1)	0.249

Basis set: I. Ema *et al.*, J. Comput. Chem., **24**, 859 (2003)
Exp: A. Kramida and W. C. Martin, *J. Phys. Chem. Ref. Data*, **26** 1185 (1997)

Wave function parametrization

- orbital coefficients $\boldsymbol{\lambda}$ are redundant $\Longrightarrow$ **bad parametrization**

Reparametrization of orbital coefficients $\boldsymbol{\lambda} \to \boldsymbol{\kappa}$ (used in MCSCF)

$$|\Psi(\mathbf{p})\rangle = \hat{J}(\boldsymbol{\alpha})\, e^{\hat{\kappa}(\boldsymbol{\kappa})} \sum_{I=1}^{N_{\text{CSF}}} c_I\, |C_I\rangle$$

- $\hat{\kappa}(\boldsymbol{\kappa})$ is the generator of **rotations in orbital space** (occ and vir):

$$\hat{\kappa}(\boldsymbol{\kappa}) = \sum_{k<l} \kappa_{kl} \left(\hat{E}_{k\to l} - \hat{E}_{l\to k} \right)$$

- $\hat{E}_{k\to l} = \hat{a}^\dagger_{k\uparrow}\hat{a}_{l\uparrow} + \hat{a}^\dagger_{k\downarrow}\hat{a}_{l\downarrow}$ is the singlet excitation operator

J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **126** 084102 (2007)
J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **128** 174101 (2008)

Wave function parametrization

- First-order derivatives with respECt to:
  - Jastrow parameters:

$$|\Psi_{\alpha_i}\rangle = \frac{\hat{J}(\boldsymbol{\alpha})}{\partial \alpha_i} \sum_{I=1}^{N_{\text{CSF}}} c_I \, |C_I\rangle$$

  - CSF parameters:

$$|\Psi_{C_I}\rangle = \hat{J}(\boldsymbol{\alpha}) \, |C_I\rangle$$

  - orbital parameters (single excitations):

$$|\Psi_{\kappa_{kl}}\rangle = \hat{J}(\boldsymbol{\alpha}) \sum_{I=1}^{N_{\text{CSF}}} c_I \hat{E}_{k\to l} \, |C_I\rangle$$

  - exponents:

$$|\Psi_{\zeta_\mu}\rangle = \hat{J}(\boldsymbol{\alpha}) \sum_{I=1}^{N_{\text{CSF}}} c_I \frac{\partial \, |C_I\rangle}{\partial \zeta_\mu}$$

J. Toulouse and C. Umrigar, *J. Chem. Phys.*, **128** 174101 (2008)

Jastrow factor

- Systematic polynomial with scaled interparticle distances
- It can have finite range (strictly zero beyond r_c) or infinite range ($r_c = \infty$)
- J_{en}, J_{ee} and J_{een} go to a constant whenever the interparticles distances are ∞

$$J = J_{en} J_{ee} J_{een} = \exp(f_{en} + f_{ee} + f_{een})$$

where

$$f_{en}(R_{i\alpha}) = \sum_{i=1}^{N_{\rm elec}} \sum_{\alpha=1}^{N_{\rm nuc}} \left[\left(\frac{a_1 R_{i\alpha}}{1 + a_2 R_{i\alpha}} + \sum_{p=2}^{N_{\rm ord}} a_{p+1} R_{i\alpha}^p \right) - \left(\frac{a_1 R_c}{1 + a_2 R_c} + \sum_{p=2}^{N_{\rm ord}} a_{p+1} R_c^p \right) \right]$$

$$f_{ee}(R_{ij}) = \sum_{i=2}^{N_{\rm elec}} \sum_{j=1}^{i-1} \left[\left(\frac{b_1 R_{ij}}{1 + b_2 R_{ij}} + \sum_{p=2}^{N_{\rm ord}} b_{p+1} R_{ij}^p \right) - \left(\frac{b_1 R_c}{1 + b_2 R_c} + \sum_{p=2}^{N_{\rm ord}} b_{p+1} R_c^p \right) \right]$$

$$f_{een}(R_{i\alpha}, R_{j\alpha}, R_{ij}) = \sum_{i=2}^{N_{\rm elec}} \sum_{j=1}^{i-1} \sum_{\alpha=1}^{N_{\rm nuc}} \sum_{p=2}^{N_{\rm ord}} \sum_{k=p-1}^{0} \sum_{l=l_{\rm max}}^{0} c_n R_{ij}^k (R_{i\alpha}^l + R_{j\alpha}^l)(R_{i\alpha} R_{j\alpha})^m, \qquad \text{where } m = \frac{p-k-l}{2}$$

- Scaling functions: $R = \frac{r}{1+kr}$ for J_{en} and J_{ee}, $\bar{R} = \frac{1}{1+kr}$ for J_{een}
- $R_c = R(r_c)$, $R_c = 1/\kappa$ for those functions for which $r_c = \infty$
- Cusp conditions easily satisfied

Wave function normalization

- Updating $\mathbf{p}_0 \to \mathbf{p}_0 + \Delta\mathbf{p}$ works for linear parameters
- It is not guaranteed to work for nonlinear parameters
- Freedom of the normalization of Ψ to alter the dependence on the nonlinear parameters

$$\begin{aligned} \bar{\Psi}(\mathbf{p}) &= N(\mathbf{p})\Psi(\mathbf{p}) \\ \bar{\Psi}(\mathbf{p}_0) &= \Psi(\mathbf{p}_0) \equiv \Psi_0 \end{aligned}$$

- $N(\mathbf{p})$ only depends on nonlinear parameters

$$\begin{aligned} \bar{\Psi}_i &= \Psi_i + N_i \Psi_0 \\ N_i &= \left(\frac{\partial N(\mathbf{p})}{\partial p_i} \right)_{\mathbf{p}=\mathbf{p}_0} \end{aligned}$$

- $N_i = 0$ for linear parameters

C. Umrigar *et al.*, *Phys. Rev. Lett.*, **98** 110201 (2007)

Wave function normalization

- Linearization of $\bar{\Psi}$

$$\bar{\Psi}_{\text{lin}} = \Psi_0 + \sum_i \Delta\bar{p}_i \bar{\Psi}_i$$

$$\Delta\bar{\mathbf{p}} = \frac{\Delta\mathbf{p}}{1 - \sum_i N_i \Delta p_i}$$

- $\Delta\bar{\mathbf{p}}$ and $\Delta\mathbf{p}$ related by an uniform scaling
- For the nonlinear parameters a fast and stable optimization is achieved if N_i are determined by imposing

$$\langle a \frac{\Psi_0}{||\Psi_0||} + (1-a) \frac{\Psi_{\text{lin}}}{||\Psi_{\text{lin}}||} | \bar{\Psi}_i \rangle = 0$$

- The choice a=1/2 guarantees finite parameter changes

C. Umrigar *et al.*, *Phys. Rev. Lett.*, **98** 110201 (2007)

Linear method over a finite MC sample

$$
\begin{aligned}
g_{L,i} &= 2\left[\left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} E_L(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_L(\mathbf{R}) \rangle \right] \\
g_{R,j} &= 2\left[\left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} E_L(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_L(\mathbf{R}) \rangle + \langle E_{L,j}(\mathbf{R}) \rangle \right] \\
H_{ij} &= \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} E_L(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} E_L(\mathbf{R}) \right\rangle \\
&- \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} E_L(\mathbf{R}) \right\rangle + \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_L(\mathbf{R}) \rangle \\
&+ \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{L,j}(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_{L,j}(\mathbf{R}) \rangle \\
S_{ij} &= \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle
\end{aligned}
$$

- $E_L(\mathbf{R}) = \hat{H}\Psi/\Psi$, $E_{L,i}$ local energy derivative with respect to p_i

Linear method: robustness

- The linear method is **equivalent to a stabilized Newton method**:

$$\begin{pmatrix} E_0 & \mathbf{g}^\mathsf{T}/2 \\ \mathbf{g}/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^\mathsf{T} \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix}$$

$$\Longleftrightarrow \begin{cases} (\mathbf{h} + 2\Delta E\, \mathbf{S}) \cdot \Delta\mathbf{p} = -\mathbf{g} \\ 2\Delta E = -\mathbf{g}^\mathsf{T} \cdot \Delta\mathbf{p} \end{cases}$$

where $\mathbf{h} = 2(\mathbf{H} - E_0\mathbf{S})$ is an approximate Hessian, and $\Delta E = E_0 - E_{\text{lin}} > 0$ is the energy stabilization

$\Longrightarrow$ **more robust than Newton method**

- In quantum chemistry, it is known as super-CI method or augmented Hessian method
- **Additional stabilization:** $H_{ij} \to H_{ij} + a\,\delta_{ij}$ where $a \geq 0$

Linear-response theory

- The generalized eigenvalue equation is written as

$$\mathbb{H} = \mathbb{\Delta}^{-1} \cdot \mathbb{L}$$

- with $\mathbb{L} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix}$ and with $\mathbb{\Delta} = \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & -\mathbf{S}^* \end{pmatrix}$
- $\mathbf{A}$ and $\mathbf{S}$ are Hermitian, $\mathbf{B}$ is symmetric $\rightarrow$ $\mathbb{L}$ and $\mathbb{\Delta}$ are Hermitian
- $\mathbb{H}$ is pseudo-Hermitian with respect to $\mathbb{L}$

$$\mathbb{H}^\dagger = \mathbb{L} \cdot \mathbb{H} \cdot \mathbb{L}^{-1}$$

- If $\mathbb{L}$ is positive definite then $\mathbb{H}$ has only real eigenvalues ω
- $\mathbb{H}$ is also pseudo-Hermitian with respect to $\mathbb{\Delta}$ but $\mathbb{\Delta}$ is not positive definite

Linear-response theory

- Expansion of the ground-state energy around the minimum

$$E_0(\mathbf{p}) = E_0(\mathbf{p}^{\min}) + \begin{pmatrix} \Delta\mathbf{p}^* & \Delta\mathbf{p} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \Delta\mathbf{p} \\ \Delta\mathbf{p}^* \end{pmatrix} + \dots$$

- where $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^{\min}$
- $\mathbb{L}$ is the Hessian of the ground-state energy at the minimum
- If the response is done around a ground state that is a properly stable minimum, $\mathbb{L}$ is positive definite and the linear-response equation gives only real excitation energies
- In practice, non-real excitation energies can be obtained if either:
  - a non-variational method is used for the ground state
  - the ground-state variational space is limited by imposing some symmetries

LR-VMC

$$
\begin{aligned}
A_{ij} &= H_{ij} - E_0 S_{ij} \\
&= \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{\mathsf{L}}(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{\mathsf{L}}(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{\mathsf{L}}(\mathbf{R}) \right\rangle \\
&\quad + \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{\mathsf{L},j}(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle E_{\mathsf{L},j}(\mathbf{R}) \right\rangle \\
&\quad - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_{\mathsf{L}}(\mathbf{R}) \rangle + 2 \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_{\mathsf{L}}(\mathbf{R}) \rangle
\end{aligned}
$$

$$
\begin{aligned}
B_{ij} &= \left\langle \frac{\overline{\Psi}_{ij}(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{H\Psi_0(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \\
&= \left\langle \frac{\Psi_{ij}(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{\mathsf{L}}(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_{ij}(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_{\mathsf{L}}(\mathbf{R}) \rangle \\
&\quad - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{\mathsf{L}}(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} E_{\mathsf{L}}(\mathbf{R}) \right\rangle \\
&\quad + 2 \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \left\langle \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle \langle E_{\mathsf{L}}(\mathbf{R}) \rangle
\end{aligned}
$$

Oscillator strengths

In the linear-response theory the oscillator strength f_n for the transition from the ground state to an excited n state is given by:

$$f_n = \frac{2}{3}\omega_n \sum_i \left[(\mathbf{X}_n + \mathbf{Y}_n)^\intercal \boldsymbol{\mu}_i \right]^2,$$

where ω_n is the excitation energy and $\boldsymbol{\mu_i} = \{\mu_i^1, \mu_i^2 ...\}$, with $\mu_i^j = \langle \overline{\Psi}_j | \mu_i | \Psi_0 \rangle$.
The j-th transition moment for the component i (x, y or z) is written as

$$\begin{aligned} \mu_i^j &= \langle \overline{\Psi}_j | \mu_i | \Psi_0 \rangle \\ &= \langle \Psi_j | \mu_i | \Psi_0 \rangle - \langle \Psi_0 | \Psi_j \rangle \langle \Psi_0 | \mu_i | \Psi_0 \rangle, \end{aligned}$$

and becomes in VMC

$$\mu_i^j = \left\langle \frac{\Psi_j r}{\Psi_0} \right\rangle - \left\langle \frac{\Psi_j}{\Psi_0} \right\rangle \langle r \rangle$$

Oscillator strengths

	CIS		LR-VMC	
f	VB1	VB2	VB1	VB2
$2s2p\ (^1P)$	0.648	0.669	0.473(2)	0.58(4)

Slater basis sets

VB1 (5,1)	VB2 (6,2,1)	VB3 (7,3, 2,1)
1s 6.285179	1s 6.285179	1s 6.335573
1s 3.455497	1s 3.455497	1s 3.466179
2s 2.774117	2s 2.679356	2s 2.686051
2s 1.192734	2s 1.183951	2s 1.171203
2s 0.824535	2s 0.790729	2s 0.953647
2p 0.986656	3s 0.479800	3s 0.922201
	2p 0.993615	3s 0.478700
	2p 0.493253	2p 1.780972
	3d 0.986170	2p 0.904739
		2p 0.567203
		3d 1.067602
		3d 0.734576
		4f 1.492274

- VB1: minimization of the sum of the HF and CISD ground-state energies
- VB2: the sum also includes the energies of the positive and negative ions
- VB3: the sum also includes the energies of a few lowest lying excited states

I. Ema *et al.*, J. Comput. Chem., **24**, 859 (2003)