

Dyson Orbitals within the fc-CVS-EOM-CCSD Framework: Theory and Application to X-ray Photoelectron Spectroscopy of Ground and Excited States.

Supplementary Information

Marta L. Vidal,^{*,†} Anna I. Krylov,^{*,‡} and Sonia Coriani^{*,†}

[†]*DTU Chemistry - Department of Chemistry, Technical University of Denmark, DK-2800,
Kongens Lyngby, Denmark*

[‡]*Department of Chemistry, University of Southern California, Los Angeles, California
90089-0482*

E-mail: malop@kemi.dtu.dk; krylov@usc.edu; soco@kemi.dtu.dk

1 Programmable expressions for Dyson orbitals within fc-CVS-EOM-CCSD

Notation

- I, J, \dots = occupied core orbital (very small number of orbitals. Core orbitals of nuclei with different Z can be handled separately, as coupling between them is negligible)
- i_v, j_v, \dots = occupied valence orbital

- generic $i, j =$ any occupied orbital (both core and valence)

In the fc-CVS-EOM-CCSD method, coupled-cluster amplitudes T and response amplitudes Λ are restricted to the valence space only (i.e., frozen-core approximation is imposed). The CVS-EOM-CCSD amplitudes R and L are constrained to the CVS domain, i.e., only the amplitudes with at least one core orbital are non-zero.

1.1 Dyson Orbitals

CCSD – CVS-EOM-IP-CCSD

$$\gamma_a^R = \langle \Phi_0(1 + \Lambda)e^{-T}|\hat{a}_a^\dagger|R_{\text{IP}}^{\text{CVS}}e^T\Phi_0\rangle = 0 \quad (\text{S1})$$

$$\gamma_{i_v}^R = \langle \Phi_0(1 + \Lambda)e^{-T}|\hat{a}_{i_v}^\dagger|R_{\text{IP}}^{\text{CVS}}e^T\Phi_0\rangle = 0 \quad (\text{S2})$$

$$\gamma_I^R = \langle \Phi_0(1 + \Lambda)e^{-T}|\hat{a}_I^\dagger|R_{\text{IP}}^{\text{CVS}}e^T\Phi_0\rangle = r_I + \sum_{k_v c} \lambda_{k_v}^c r_{Ik_v}^c \quad (\text{S3})$$

$$\gamma_a^L = \langle \Phi_0L_{\text{IP}}^{\text{CVS}}e^{-T}|\hat{a}_a|e^T\Phi_0\rangle = 0 \quad (\text{S4})$$

$$\gamma_{i_v}^L = \langle \Phi_0L_{\text{IP}}^{\text{CVS}}e^{-T}|\hat{a}_{i_v}|e^T\Phi_0\rangle = 0 \quad (\text{S5})$$

$$\gamma_I^L = \langle \Phi_0L_{\text{IP}}^{\text{CVS}}e^{-T}|\hat{a}_I|e^T\Phi_0\rangle = l_I \quad (\text{S6})$$

CVS-EOM-EE-CCSD – CVS-EOM-IP-CCSD¹

$$\gamma_a^R = \langle \Phi_0 L_{EE}^{\text{CVS}} e^{-T} | \hat{a}_a^\dagger | R_{IP}^{\text{CVS}} e^T \Phi_0 \rangle = \sum_K l_K^a r_K + \frac{1}{2} \sum_{KLc} l_{LK}^{ca} r_{KL}^c + \sum_{Kl_v c} l_{l_v K}^{ca} r_{Kl_v}^c \quad (\text{S7})$$

$$\begin{aligned} \gamma_{i_v}^R &= \langle \Phi_0 L_{EE}^{\text{CVS}} e^{-T} | \hat{a}_{i_v}^\dagger | R_{IP}^{\text{CVS}} e^T \Phi_0 \rangle \\ &= \sum_{Kc} l_K^c r_{i_v K}^c - \sum_c \gamma_c^{R, \text{virt}} t_{i_v}^c - \frac{1}{2} \sum_K r_K \sum_{l_v cd} l_{l_v K}^{dc} t_{l_v i_v}^{dc} \end{aligned} \quad (\text{S8})$$

$$\gamma_I^R = \langle \Phi_0 L_{EE}^{\text{CVS}} e^{-T} | \hat{a}_I^\dagger | R_{IP}^{\text{CVS}} e^T \Phi_0 \rangle = \sum_{Kc} l_K^c r_{IK}^c \quad (\text{S9})$$

$$\begin{aligned} \gamma_a^L &= \langle \Phi_0 L_{IP}^{\text{CVS}} e^{-T} | \hat{a}_a | R_{EE}^{\text{CVS}} e^T \Phi_0 \rangle \\ &= \sum_K l_K r_K^a + \frac{1}{2} \sum_{KLc} l_{KL}^c r_{KL}^{ac} + \sum_{Kl_v c} l_{Kl_v}^c r_{Kl_v}^{ac} + \sum_{k_v Lc} l_{k_v L}^c t_{k_v}^a r_L^c \end{aligned} \quad (\text{S10})$$

$$\gamma_{i_v}^L = \langle \Phi_0 L_{IP}^{\text{CVS}} e^{-T} | \hat{a}_{i_v} | R_{EE}^{\text{CVS}} e^T \Phi_0 \rangle = \sum_{Kc} l_{i_v K}^c r_K^c \quad (\text{S11})$$

$$\gamma_I^L = \langle \Phi_0 L_{IP}^{\text{CVS}} e^{-T} | \hat{a}_I | R_{EE}^{\text{CVS}} e^T \Phi_0 \rangle = l_I r_0 + \sum_{Kc} l_{IK}^c r_K^c \quad (\text{S12})$$

¹The superscript *virt* in S9 emphasizes the use of the virtual block (S7)

EOM-EE-CCSD – CVS-EOM-IP-CCSD

$$\gamma_a^R = \langle \Phi_0 L_{EE} e^{-T} | \hat{a}_a^\dagger | R_{IP}^{CVS} e^T \Phi_0 \rangle = 0 \quad (\text{S13})$$

$$\gamma_{i_v}^R = \langle \Phi_0 L_{EE} e^{-T} | \hat{a}_{i_v}^\dagger | R_{IP}^{CVS} e^T \Phi_0 \rangle = 0 \quad (\text{S14})$$

$$\gamma_I^R = \langle \Phi_0 L_{EE} e^{-T} | \hat{a}_I^\dagger | R_{IP}^{CVS} e^T \Phi_0 \rangle = \sum_{k_v c} l_{k_v}^c r_{I k_v}^c \quad (\text{S15})$$

$$\gamma_a^L = \langle \Phi_0 L_{IP}^{CVS} e^{-T} | \hat{a}_a | R_{EE} e^T \Phi_0 \rangle = 0 \quad (\text{S16})$$

$$\gamma_{i_v}^L = \langle \Phi_0 L_{IP}^{CVS} e^{-T} | \hat{a}_{i_v} | R_{EE} e^T \Phi_0 \rangle = 0 \quad (\text{S17})$$

$$\gamma_I^L = \langle \Phi_0 L_{IP}^{CVS} e^{-T} | \hat{a}_I | R_{EE} e^T \Phi_0 \rangle = l_I r_0 + \sum_{k_v c} l_{I k_v}^c r_{k_v}^c \quad (\text{S18})$$

CVS-EOM-EE-CCSD – EOM-IP-CCSD ²

$$\gamma_a^R = \langle \Phi_0 L_{EE}^{CVS} e^{-T} | \hat{a}_a^\dagger | R_{IP} e^T \Phi_0 \rangle = 0 \quad (\text{S19})$$

$$\gamma_{i_v}^R = \langle \Phi_0 L_{EE}^{CVS} e^{-T} | \hat{a}_{i_v}^\dagger | R_{IP} e^T \Phi_0 \rangle = 0 \quad (\text{S20})$$

$$\gamma_I^R = \langle \Phi_0 L_{EE}^{CVS} e^{-T} | \hat{a}_I^\dagger | R_{IP} e^T \Phi_0 \rangle = 0 \quad (\text{S21})$$

$$\gamma_a^L = \langle \Phi_0 L_{IP} e^{-T} | \hat{a}_a | R_{EE}^{CVS} e^T \Phi_0 \rangle = r_0 \gamma_a^{L, \text{CC-IP}} \quad (\text{S22})$$

$$\gamma_{i_v}^L = \langle \Phi_0 L_{IP} e^{-T} | \hat{a}_{i_v} | R_{EE}^{CVS} e^T \Phi_0 \rangle = l_{i_v} r_0 \quad (\text{S23})$$

$$\gamma_I^L = \langle \Phi_0 L_{IP} e^{-T} | \hat{a}_I | R_{EE} e^T \Phi_0 \rangle = 0 \quad (\text{S24})$$

²The superscript CC-IP in S22 refers to the (non-CVS) EOMIP block:
 $\gamma_a^L = \sum_{k_v} t_{k_v}^a l_{k_v} + \frac{1}{2} \sum_{k_v l_v c} t_{k_v l_v}^{ac} l_{k_v l_v}^c$

2 TR-XPS with the ground-state reference

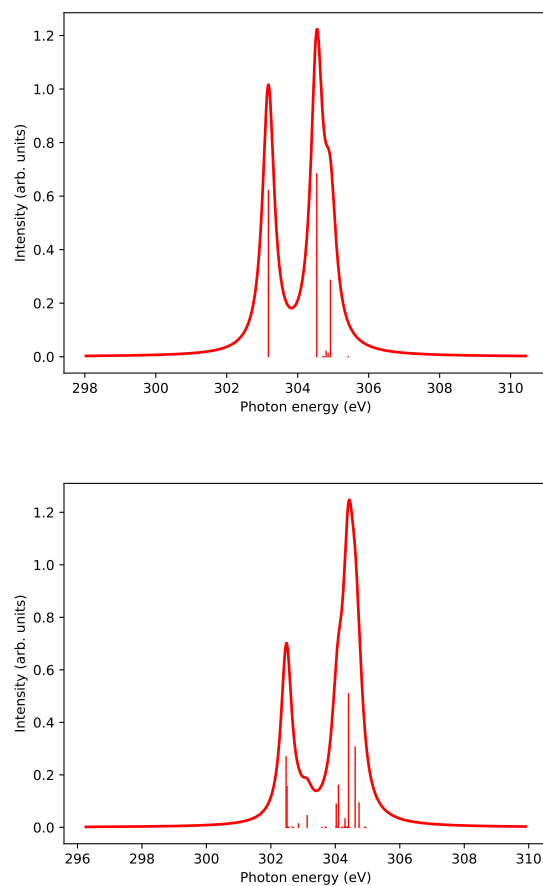


Figure S1: Uracil. Carbon K-edge fc-CVS-EOM-CCSD/6-311(2+,+)G** (uncontracted) X-ray photoelectron spectra of the first excited state (S_1) and second excited state (S_2) obtained by convolution of the computed ionization energies and Dyson norms with a Lorentzian function (FWHM = 0.4 eV).

Table S1: Uracil. Carbon K-edge fc-CVS-EOM-CCSD/6-311(2+,+)G** (uncontracted) ionization energies (IE, eV) and squared norms of Dyson orbitals for the first (S_1), and second (S_2) excited states. IEs are computed at the Franck-Condon geometry.

S_1		S_2	
IE	$ \phi^{\text{Dyson}} ^2$	IE	$ \phi^{\text{Dyson}} ^2$
303.03	0.000	301.26	0.000
303.17	0.620	301.85	0.000
304.20	0.000	302.47	0.268
304.54	0.683	302.50	0.155
304.71	0.000	302.56	0.000
304.80	0.020	302.86	0.013
304.80	0.000	303.02	0.000
304.86	0.013	303.13	0.044
304.92	0.285	303.59	0.000
305.43	0.000	304.03	0.088
		304.10	0.159
		304.13	0.000
		304.30	0.032
		304.41	0.508
		304.61	0.306
		304.74	0.093

2.1 Comparison of IEs computed with fc-CVS-EOM-CCSD and within Hartree-Fock/Koopmans approximation

Table S2: Adenine. Carbon K-edge ionization energies (IE, eV) computed with Koopmans theorem and with fc-CVS-EOM-CCSD with the 6-311(2+,+)G** (uncontracted on C) basis set.

#	Koopmans	fc-CVS-EOM-IP-CCSD	Exp. ¹
1	306.32	291.50	291.0
2	307.87	292.94	
3	308.09	293.11	292.5
4	308.25	293.38	
5	308.66	293.76	

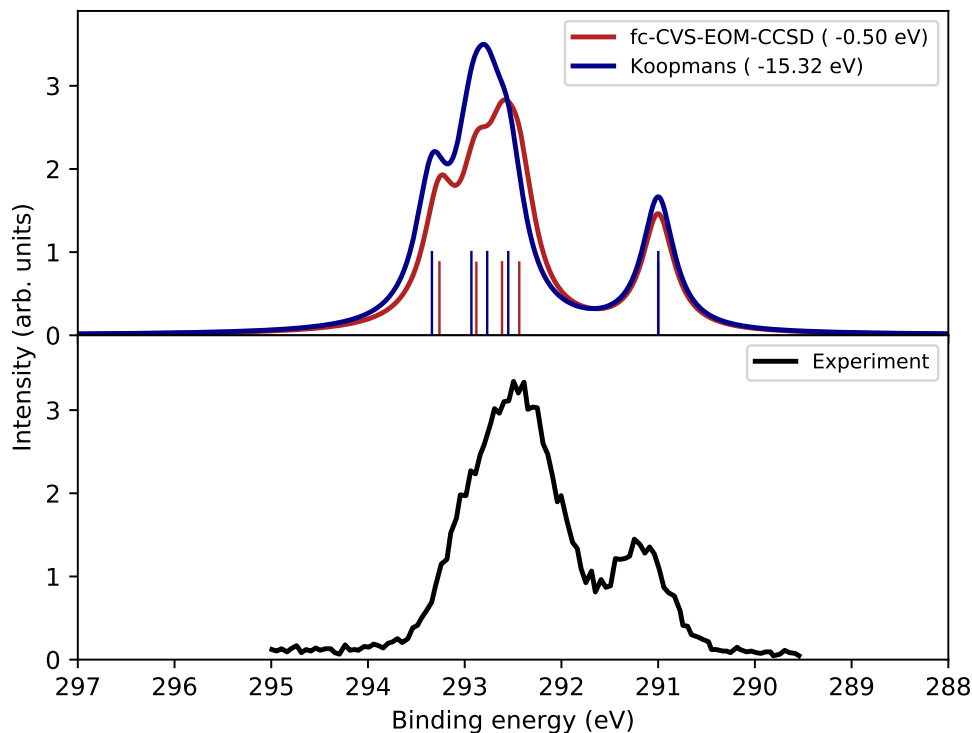


Figure S2: Adenine. Carbon K-edge X-ray photoelectron (XPS) spectra obtained at the fc-CVS-EOM-IP-CCSD/6-311(2+,+)G** (uncontracted on C) level of theory and using Koopmans theorem, by convolution of the computed ionization energies and Dyson norms with a Lorentzian function (FWHM = 0.4 eV). The uniform shift applied is given in parenthesis in the legend. The shift was determined with respect to the first experimental peak position, reported to be at 291.0 eV. The experimental spectrum was digitized from Ref. 1.

2.2 Basis Set Comparison

Table S3: Adenine. Carbon K-edge ionization energies (IE, eV) obtained with different basis sets.

#	6-311(2+,+) G^{**} (uncontracted on C)	6-311(2+,+) G^{**} (uncontracted)	aug-cc-pVTZ /aug-cc-pVDZ	aug-cc-pCVTZ /aug-cc-pVDZ	Exp. ¹
1	291.50	291.50	291.49	291.27	291.0
2	292.94	292.93	292.97	292.73	
3	293.11	293.11	293.10	292.89	
4	293.38	293.38	293.41	293.16	292.5
5	293.76	293.76	293.77	293.56	

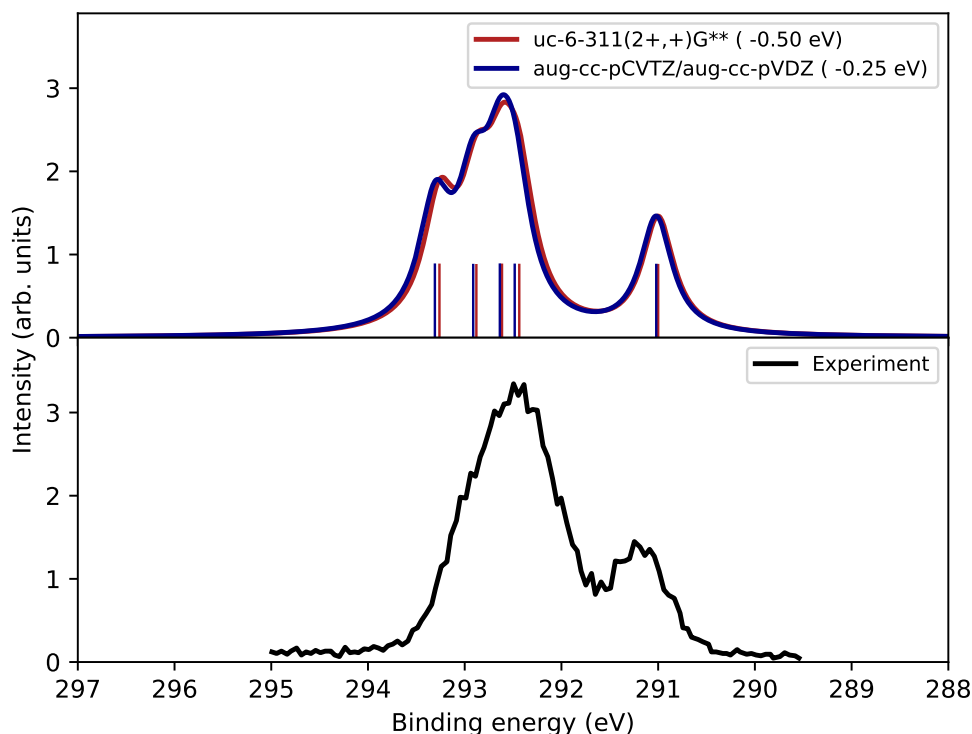


Figure S3: Adenine. Carbon K-edge fc-CVS-EOM-CCSD X-ray photoelectron (XPS) spectra obtained using different basis sets and by convolution of the computed ionization energies and Dyson norms with a Lorentzian function (FWHM = 0.4 eV). The uniform shift applied is given in parenthesis in the legend. The shift was determined with respect to the first experimental peak position, reported to be at 291.0 eV. The experimental spectrum was digitized from Ref. 1.

3 Relevant Cartesian geometries

3.1 Adenine

Table S4: Optimized geometry of adenine at the MP2/cc-pVTZ level of theory (Å).

Atom	x	y	z
N	1.663948	-1.127531	0.000000
N	-0.643862	-1.804088	0.000000
N	-0.636159	1.773662	0.000000
N	-2.146840	0.119893	0.000000
N	2.353671	1.084678	0.000000
C	0.668787	-2.022720	0.000000
C	-0.919981	-0.499328	0.000000
C	0.000000	0.546991	0.000000
C	1.356490	0.175959	0.000000
C	-1.907009	1.474843	0.000000
H	-3.039931	-0.341646	0.000000
H	-2.712908	2.189994	0.000000
H	0.978346	-3.060960	0.000000
H	2.144452	2.065199	0.000000
H	3.305020	0.766647	0.000000

3.2 Uracil

Table S5: Optimized geometry of the ground state of uracil at the MP2/cc-pVTZ level of theory (Å).

Atom	x	y	z
C	0.052842	-1.701332	0.000000
H	-2.016793	-1.512231	0.000000
H	-0.025494	-2.778299	0.000000
H	-0.023568	1.988544	0.000000
C	-1.232948	0.358893	0.000000
O	-2.294123	0.950188	0.000000
C	1.274970	0.391620	0.000000
O	2.275873	1.086244	0.000000
N	0.000000	0.978006	0.000000
N	-1.137535	-1.023018	0.000000
C	1.239239	-1.061325	0.000000
H	2.169983	-1.601520	0.000000

Table S6: Optimized geometry of the first valence excited state of uracil ($S_1 n\pi^*$) at the CCSD/aug-cc-pVDZ level of theory (Å).

Atom	x	y	z
C	-0.0042112405	-1.7549930536	0.0000000000
H	-2.0949460216	-1.4184138214	0.0000000000
H	-0.1300971688	-2.8358204634	0.0000000000
H	0.0374086733	2.0243048927	0.0000000000
C	-1.2210197500	0.4010705948	0.0000000000
O	-2.2535781024	1.0567627444	0.0000000000
C	1.2399442445	0.2839850904	0.0000000000
O	2.3361887279	1.0691923037	0.0000000000
N	0.0349480223	1.0125648826	0.0000000000
N	-1.1853832587	-0.9770870742	0.0000000000
C	1.2251302704	-1.1165150648	0.0000000000
H	2.1596039787	-1.6767971209	0.0000000000

Table S7: Optimized geometry of the second valence excited state of uracil ($S_2 \pi\pi^*$) at the CCSD/aug-cc-pVDZ level of theory (Å).

Atom	x	y	z
C	-0.0581949473	-1.7976206857	0.0000000000
H	-2.0937197909	-1.3962500862	0.0000000000
H	-0.1994763750	-2.8770618511	0.0000000000
H	0.0643245546	2.0452320710	0.0000000000
C	-1.1889222032	0.4314192107	0.0000000000
O	-2.2890290202	0.9759697793	0.0000000000
C	1.3011427138	0.3335152409	0.0000000000
O	2.3251617259	1.0557727555	0.0000000000
N	0.0308489641	1.0316683625	0.0000000000
N	-1.1550971105	-0.9985052937	0.0000000000
C	1.2377799195	-1.0849263853	0.0000000000
H	2.1691699442	-1.6509592078	0.0000000000

References

- (1) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K.; Trofimov, A.; Gromov, E.; Zaytseva, I.; Schirmer, J. A theoretical and experimental study of the near edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron spectra (XPS) of nucleobases: Thymine and adenine. *Chem. Phys.* **2008**, *347*, 360–375.