Supporting Information

Steady and time-resolved photoelectron spectra based on nuclear ensembles

Wilmer Arbelo-González, Rachel Crespo-Otero, Mario Barbatti

Table of Contents

SI-1. Dyson orbitals

The Dyson orbitals (DO) associated with a particular $I \rightarrow F$ transition is a single electron wavefunction containing information on where the ejected electron was removed from. This section is devoted to find explicit expressions to calculate the DOs. To do that, let us first define the vectors $\vec{\mathbf{r}}_N = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ and $\vec{\mathbf{r}}_{N-1} = (\mathbf{r}_1, \ldots, \mathbf{r}_{N-1})$, so that the DO is written as

$$\psi_{IF}^d(\mathbf{r}_N) = \sqrt{N} \int d\vec{\mathbf{r}}_{N-1} \, \tilde{\Phi}_F(\vec{\mathbf{r}}_{N-1}) \, \Phi_I(\vec{\mathbf{r}}_N), \tag{1}$$

where N is the number of electrons in the I state.

Consider now that Φ_I and $\tilde{\Phi}_F$ can be represented as a linear combination of Slater determinants, *i.e.*,

$$\Phi_I(\vec{\mathbf{r}}_N) = \sum_{n=0}^{N_{max}} c_n \Psi_n(\vec{\mathbf{r}}_N) \tag{2}$$

and

$$\tilde{\Phi}_F(\vec{\mathbf{r}}_{N-1}) = \sum_{m=0}^{M_{max}} \tilde{c}_m \tilde{\Psi}_m(\vec{\mathbf{r}}_{N-1}), \tag{3}$$

where N_{max} and M_{max} are the maximum number of terms used in the expansions. It is shown in the Section SI-1.1 of this Supporting Information, that each Slater determinant in Eq. (2) can be rewritten as

$$\Psi_n = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} (-1)^{N+j} \Psi_n^j (\vec{\mathbf{r}}_{N-1}) \chi_q^{(j)}(\mathbf{r}_N), \tag{4}$$

where $\Psi_n^j (\vec{\mathbf{r}}_{N-1})$ is the $(N-1) \times (N-1)$ Slater determinant associated with the minor that results from removing the Nth row and the jth column out of $\Psi_n(\vec{\mathbf{r}}_N)$. The superscript in the spin-orbital $\chi_q^{(j)}(\mathbf{r}_N)$ in Eq. (4) denotes the column occupied in $\Psi_n(\vec{\mathbf{r}}_N)$.

By inserting Eq. (3) and Eq. (2) into Eq. (1), with the Ψ_n given by Eq. (4), the DO can be rewritten as

$$\psi_{IF}^d (\mathbf{r}_N) = \sum_{n=0}^{N_{max}} \sum_{j=1}^{N} d_{nj}^q \chi_q^{(j)}(\mathbf{r}_N), \tag{5}$$

with

$$d_{nj}^q = (-1)^{N+j} c_n \sum_{m=0}^{M_{max}} \tilde{c}_m \langle \tilde{\Psi}_m | \Psi_n^j \rangle. \tag{6}$$

In Eq. (6), the integration is over $\vec{\mathbf{r}}_{N-1}$ and the q superscript denotes the actual index of the spin-orbital removed from the jth column.

Regarding this, the spin-orbital $\chi_q^{(j)}$, removed from the jth column out of each Slater determinant Ψ_n, can be in principle any of the N_{bf} spin-orbitals conforming the basis set used to describe the electronic state, i.e. $q \in \{1, \ldots, N_{bf}\}$. Therefore, by grouping the d_{nj}^q coefficients corresponding to the same spin-orbitals, Eq. (5) can be conveniently rearranged as

$$\psi_{IF}^d (\mathbf{r}_N) = \sum_{s=1}^{N_{bf}} b_s \chi_s (\mathbf{r}_N), \tag{7}$$

with

$$b_s = \sum_{n=0}^{N_{max}} \sum_{j=1}^{N} d_{nj}^q \delta_{qs}, \tag{8}$$

where δ_{qs} is the Kronecker delta function. Using Eq. (7), the norm of the DO can be easily computed as

$$\left\| \psi_{IF}^d \right\| = \left(\sum_{s=1}^{N_{bf}} b_s^2 \right)^{1/2}, \tag{9}$$

which, in general, is not equal to one.

Eq. (7) also illustrates the fact that the DOs are formed from a linear combination of spin-orbitals of the initial state. The contribution of each spin-orbital to the DO will depend on (i) the expansion coefficients c_n and $\tilde{c}_m$ of the electronic wavefunctions and (ii) the overlaps between single Slater determinants of the initial and final states. The first quantity is specific to the method used to solve the electronic problem and it is discussed for the particular case of TDDFT in the main paper. The second quantity can be computed as follows.

Suppose $\tilde{\Psi}_m = |\tilde{\chi}_1 \cdots \tilde{\chi}_p \cdots \tilde{\chi}_b\rangle$ is a given Slater determinant of the F state and $\Psi_n^j = |\chi_1 \cdots \chi_q \cdots \chi_a\rangle$ one associated with a minor of the I state, both of range $(N-1)\times(N-1)$. To simplify the notation, the superscript denoting the column has been omitted. Remember that, as discussed in the introduction, within the Δ-methods framework, $\{\chi_q\}$ and $\{\tilde{\chi}_p\}$ are determined by two independent calculations for the N and $N-1$ systems, and they are not orthogonal to each other. It is possible to show that the overlap between the Slater determinants in Eq. (6) can be written as the following determinant

$$\langle \tilde{\Psi}_m | \Psi_n^j \rangle = \begin{vmatrix} \langle \tilde{\chi}_1 | \chi_1 \rangle & \cdots & \langle \tilde{\chi}_1 | \chi_q \rangle & \cdots & \langle \tilde{\chi}_1 | \chi_a \rangle \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \langle \tilde{\chi}_p | \chi_1 \rangle & \cdots & \langle \tilde{\chi}_p | \chi_q \rangle & \cdots & \langle \tilde{\chi}_p | \chi_a \rangle \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \langle \tilde{\chi}_b | \chi_1 \rangle & \cdots & \langle \tilde{\chi}_b | \chi_q \rangle & \cdots & \langle \tilde{\chi}_b | \chi_a \rangle \end{vmatrix}. \tag{10}$$

To compute each element in the previous equation, we first represent the spin-orbitals in the atomic orbitals basis set, *i.e.*,

$$\tilde{\chi}_p = \sum_{u=1}^{N_{bf}} \tilde{a}_{up} \phi_u \tag{11}$$

and

$$\chi_q = \sum_{v=1}^{N_{bf}} a_{vq} \phi_v. \tag{12}$$

Notice that as the ionization is very fast as compared with the nuclear motions, the nuclear positions are considered fixed while the electron is ejected and therefore the same atomic basis set $\{\phi\}$ has been used in the expansions. Thus, by using Eqs. (11) and (12),

$$\langle \tilde{\chi}_p | \chi_q \rangle = \sum_{u,v=1}^{N_{bf}} \tilde{a}_{up} a_{vq} S_{uv}, \tag{13}$$

where $S_{uv} = \langle \phi_u | \phi_v \rangle$ is the overlap matrix between the atomic orbitals, a standard output when computing the electronic states.

To recap, the calculation of the DOs requires first to solve the electronic problem and represent the electronic wavefunctions of the initial and final states as a linear combination of Slater determinants. Then, with (i) the expansion coefficients c_n and $\tilde{c}_m$, (ii) the spin-orbitals $\{\chi_p\}$ and $\{\tilde{\chi}_q\}$ in the atomic orbitals representation, and (iii) the overlap matrix S_{uv}, the DOs and their norms can be computed by means of Eqs. (5)-(9).

To reduce computational costs, d_{nj}^q terms with expansion coefficient c_n or $\tilde{c}_m$ smaller than ε_{cis} can be neglected. In all results discussed here, we have adopted $\varepsilon_{cis} = 0.01$.

SI-1.1 Demonstration of Eq. (4)

A typical $N \times N$ Slater determinant in Eq. (35) for a particular set of spin-orbitals, say for instance, $\{\chi_1 \cdots \chi_q \cdots \chi_a\}$, reads

$$\Psi_n = \frac{1}{\sqrt{N!}} \sum_{j=1}^{N} (-1)^{N+j} \mid \chi_1^{(1)} \cdots \chi_a^{(N-1)} \rangle \chi_q^{(j)}(\mathbf{r}_N), \tag{14}$$

where a superscript in the spin-orbitals has been intentionally added to denote the column they occupy.

This generic determinant in Eq. (14) can be expanded over the Nth row (Laplace's formula), leading to

$$\Psi_n = \frac{1}{\sqrt{N!}} \sum_{j=1}^{N} (-1)^{N+j} \mid \chi_1^{(1)} \cdots \chi_a^{(N-1)} \rangle \chi_q^{(j)}(\mathbf{r}_N), \tag{15}$$

where $\mid \chi_1^{(1)} \cdots \chi_a^{(N-1)} \rangle$ is the $(N-1) \times (N-1)$ minor that results from removing the Nth row and the jth column (initially occupied by the spin-orbital $\chi_q^{(j)}$) out of the determinant in Eq. (14). Denoting the Slater determinant associated with $\mid \chi_1^{(1)} \cdots \chi_a^{(N-1)} \rangle$ as

$$\Psi_n^j(\vec{\mathbf{r}}_{N-1}) = \frac{1}{\sqrt{(N-1)!}} \mid \chi_1^{(1)} \cdots \chi_a^{(N-1)} \rangle, \tag{16}$$

Eq. (15) can be rewritten as

$$\Psi_n = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} (-1)^{N+j} \Psi_n^j(\vec{\mathbf{r}}_{N-1}) \chi_q^{(j)}(\mathbf{r}_N). \tag{17}$$

SI-2. Ionization rules

For practical purposes, it is useful to have a simple way to predict whether or not the ionization is allowed, without explicit calculation of Dyson orbitals and transition dipoles. For decades, such analysis has been carried out based on orbitals analysis, which distinguishes between single- and double-electron processes leading to, respectively, allowed and forbidden single ionization.[1] In particular, the double-electron forbidden processes can borrow intensity from allowed transitions by configuration interaction giving rise to weak shake-up bands.[2-3] In this section, we will briefly review the theory associated to these allowed and forbidden processes in terms of the developments presented so far.

To do so, let us consider a given $I \to F$ photoionization, where Φ_I and $\tilde{\Phi}_F$ are dominated by only one electronic configuration each, not necessarily the one corresponding to the ground state. For instance, assume r and s are the indexes of the Slater determinants respectively dominating the I and F states. Thus, the expansion coefficients $c_n = \delta_{rn}$ and $\tilde{c}_m = \delta_{sm}$. By using Eqs. (5) and (6), the DO reduces to

$$\psi_{IF}^d = \sum_{j=1}^{N} (-1)^{N+j} \langle \tilde{\Psi}_r \mid \Psi_s^j \rangle \chi_q^{(j)}. \tag{18}$$

Notice that Ψ_s^j, of range $(N-1)\times(N-1)$, is obtained by removing the spin-orbital at column j from Ψ_s. In addition, let us assume in a first step that the Koopmans approximation is valid and the set of spin-orbitals used to describe the I and F electronic states are strictly the same.

Now, suppose that the particular electronic arrangements corresponding to Ψ_r and $\tilde{\Psi}_s$ differ by only one spin-orbital. For instance, consider the case $\Psi_r = |\chi_1^{(1)} \cdots \chi_k^{(i)} \cdots \chi_a^{(N)}\rangle$ and $\tilde{\Psi}_s = |\chi_1^{(1)} \cdots \chi_a^{(N-1)}\rangle$, with χ_k located at column i being the different one. According to Eq. (10), unless χ_k is removed from Ψ_r, the whole column with the overlaps of χ_k with each of the spin-orbitals of $\tilde{\Psi}_s$ will be zero, making zero the determinant, the DO, and, therefore, the transition dipole. This means analytically that $\langle\tilde{\Psi}_r | \Psi_s^j\rangle = \delta_{ij}$, which reduces Eq. (18) to the simple form

$$\psi_{IF}^d = (-1)^{N+i} \chi_k, \tag{19}$$

with norm $\left\|\psi_{IF}^d\right\| \equiv \left\|\chi_k\right\| = 1$. This well-known result[4] can be formulated in what it will be called in the following as the *first ionization rule*.

First ionization rule: *when the electronic configuration describing the initial state differs by only one spin-orbital from the electronic configuration of the final state, the ionization is possible and the Dyson orbital matches, except possibly by a sign, this different spin-orbital.*

Consider now the situation when Ψ_r differs by two spin-orbitals with respect to $\tilde{\Psi}_s$. For instance, assume that $\Psi_r = |\chi_1^{(1)} \cdots \chi_k^{(i)} \chi_l^{(i+1)} \cdots \chi_a^{(N)}\rangle$ and $\tilde{\Psi}_r = |\chi_1^{(1)} \cdots \chi_a^{(N-1)}\rangle$, with χ_k and χ_l being the different ones. In this case, following the same reasoning as before, it does not matter which one of the different spin-orbitals we remove, there will always be a whole column full of zeros in Eq. (10), corresponding to the overlaps of the other one with all the spin-orbitals of $\tilde{\Psi}_s$. As a result, the DO and, therefore, the transition dipole will always vanish. By the same argument, if there are more than two different spin-orbitals between Ψ_r and $\tilde{\Psi}_s$, the transition dipole will also vanish. This can be formulated as the *second ionization rule*.

Second ionization rule: *when the electronic configuration describing the initial state differs by two or more spin-orbitals from the electronic configuration of the final state, the ionization is not allowed.*

Naturally, these rules are rigorously valid only within the Koopmans approximation. Many-body effects, such as orbital relaxation and correlation shifts after hole creation, will always cause changes in the intensities.[2] However, in many practical problems, a Koopmans' analysis gives a fair description of the electronic wavefunction. Thus, these rules can still be applied for a qualitative evaluation. Let us examine first imidazole ($N = 36$ electrons).

According to the DFT results for the $0 \to 0$ transition, the electronic ground state wavefunctions of the neutral molecule and the cation are $\Phi_0 = |1\,\overline{1}\cdots17\,\overline{17}\,18\,\overline{18}\rangle$ and $\tilde{\Phi}_0 = |1\,\overline{1}\cdots17\,\overline{17}\,18\rangle$, respectively. They are both represented in the left panel of Figure S1. Here q ($\overline{q}$) denotes the qth α (β) orbital. As can be noticed, the electronic configurations of both species differ by only one spin-orbital, the $\overline{\chi}_{18}$. Thus, according to the first ionization rule, the ionization is allowed and $\psi_{00}^d(\mathbf{R}_0) \approx \overline{\chi}_{18}$. The accurate calculation of the DO shows that 99% of the DO comes, in fact, from the $\overline{\chi}_{18}$ spin-orbital.

For the $0 \to 1$ transition of imidazole, represented in the right panel of Figure S1, the TDDFT calculations show that $\tilde{\Phi}_1 \approx |1\,\bar{1}\cdots17\,\overline{18}\,18\rangle$. According to the first ionization rule, the ionization is allowed and now $\psi_{01}^d(\mathbf{R}_0) \approx \bar{\chi}_{17}$. Once again, the accurate calculation of the DO confirms that.

Up to $F = 3$, the cationic electronic wavefunctions are mostly dominated by one Slater determinant and a similar analysis is possible. For $F \geq 4$ and especially for the higher states, several electronic configurations are involved in the wavefunctions and the analysis has to be performed individually for each Slater determinant. Even so, for all transitions, the ionization rules are able to predict qualitatively the norm and composition of the DOs.
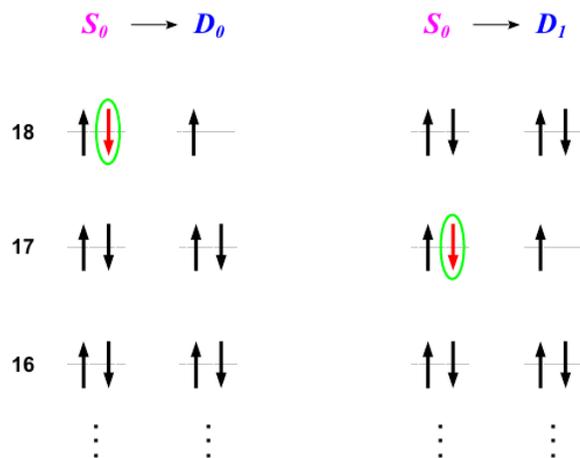

Figure S1 - Electronic configurations of the neutral molecule and the cation for the $0 \to 0$ (left) and $0 \to 1$ (right) photoionizations of imidazole. In red, the orbitals of the neutral molecule that do not match any of the orbitals of the cation. When enclosed with a green line, they correspond to the DO.

For adenine ($N = 70$ electrons), with electronic ground state wavefunction $\Phi_0 = |1\,\bar{1}\cdots34\,\overline{34}\,35\,\overline{35}\rangle$, the ionization rules are also able to (qualitatively) predict the norm and shape of the DOs for each $0 \to F$ transition. Let us show that with two illustrative examples, represented in Figure S2. The first one is the $0 \to 3$ transition (left panel), for which $\tilde{\Phi}_3 \approx |1\,\bar{1}\cdots31\,\overline{35}\,32\,\overline{32}\cdots34\,\overline{34}\,35\rangle$. As the only different spin-orbital between $\tilde{\Phi}_3$ and Φ_0 is $\bar{\chi}_{32}$, invoking the first ionization rule, $\psi_{03}^d(\mathbf{R}_0) \approx \bar{\chi}_{32}$. This is confirmed by the accurate calculation of the DO, which shows that the DO has a 99% contribution from the $\bar{\chi}_{32}$ spin-orbital.

The second example is the $0 \to 8$ transition (right panel in Figure S2). For $F = 8$, the major contribution in $\tilde{\Phi}_8$ comes from the Slater determinant $\tilde{\Psi} = |1\,\bar{1}\cdots33\,\overline{33}\,36\,\overline{34}\,35\rangle$. As χ_{34} and $\bar{\chi}_{35}$ in Φ_0 do not match any of the spin-orbitals of $\tilde{\Psi}$, according to the second ionization rule, the ionization is not allowed. The accurate calculation of the DO shows that $\left\|\psi_{08}^d(\mathbf{R}_0)\right\| \approx 0.1$, which leads to an (approximate) ionization probability of $P_{08} \approx \left\|\psi_{08}^d(\mathbf{R}_0)\right\|^2 \approx 0.01$, in practice a very weak shake-up band.
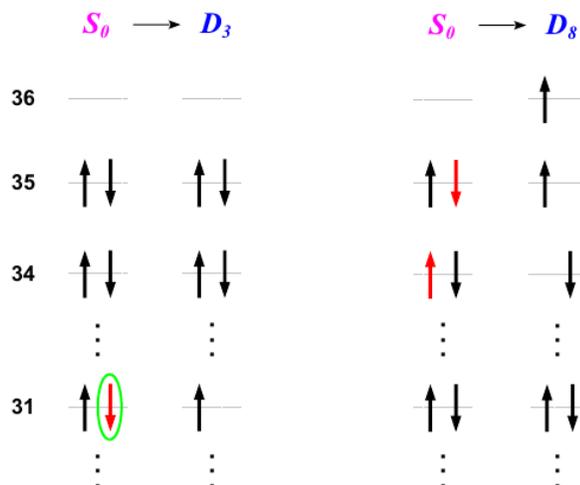
Figure S2 - Electronic configurations of the neutral molecule and the cation for the 0 → 3 (left) and 0 → 8 (right) photoionizations of adenine. In red, the orbitals of the neutral molecule that do not match any of the orbitals of the cation. When enclosed with a green line, they correspond to the DO.

SI-3. Computational Times

Here we provide some guidelines to estimate the costs for simulation of steady and time-resolved photoelectron spectra using the methods discussed in the paper.

For the steady spectrum, the major cost is to compute the energies for each point of the ensemble. In our case, 500 geometries, this is the time to run 500 independent single-point TDDFT energy calculations. The time to compute the DOs is marginal compared to that. The time to compute the cross sections is significant, but still shorter than of a single point calculation.

In the case of the time-resolved spectrum, most of time is spent on the dynamics itself. This time is about the number of trajectories times the number of time steps in each trajectory times the time for a single point TDDFT calculation of energies and gradient. In our case, this amounts to 100,000 single point calculations. The cost of the spectrum is the cost of a steady spectrum times the number of time windows in which the dynamics is split. In our case, we split 500 fs in intervals of 20 intervals of 25 fs. Therefore, the cost of the time-resolved spectrum was about 20 times that of the steady spectrum.

Based on the discussion above, concerning time, the system size limitation is basically controlled by the electronic structure calculations. If for a certain molecule we can afford to compute about 500 single points, then the steady spectrum is feasible. If we can afford 100,000 single points, the time-resolved spectrum is feasible.

Finally note that most of steps above are independent calculations that can be trivially parallelized and distributed over as many machines as available.

References

1.   Kluge, G.; Scholz, M., Chemistry of excited states. XI. Interpretation of photoelectron spectra by final state single CI method within semiempirical formalisms. *Int. J. Quantum Chem.* **1981**, *20*, 669-682.

2.   Cederbaum, L. S.; Domcke, W.; Schirmer, J.; Niessen, W. v., Many-Body Effects in Valence and Core Photoionization of Molecules. *Phys. Scr.* **1980**, *21*, 481.

3.   Ortiz, J. V., Orbital and shakeup operator renormalizations in electron propagator theory. *J. Chem. Phys.* **1998**, *109*, 5741-5746.

4.   Oana, C. M.; Krylov, A. I., Dyson orbitals for ionization from the ground and electronically excited states within equation-of-motion coupled-cluster formalism: Theory, implementation, and examples. *J. Chem. Phys.* **2007**, *127*, 234106.