Quantum Monte Carlo Treatment of the Charge Transfer and Diradical Electronic Character in a Retinal Chromophore Minimal Model

Andrea Zen,[*,†,‡] Emanuele Coccia,[*,¶,§] Samer Gozem,[‖] Massimo Olivucci,[*,‖,⊥] and Leonardo Guidoni[*,#]

Dipartimento di Fisica, Sapienza-Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy, London Centre for Nanotechnology, University College London, London WC1E 6BT, United Kingdom, Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, Via Vetoio 2, 67010 L'Aquila, Italy, Sorbonne Universités, UPMC Univ Paris 06 & CNRS, Laboratoire de Chimie Théorique, France, Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403, United States, Dipartimento di Chimica, Universitá di Siena, via De Gasperi 2, I-53100 Siena, Italy, and Dipartimento di Scienze Fisiche e Chimiche, Università degli studi dell'Aquila, Via Vetoio 2, 67010 L'Aquila, Italy

E-mail: a.zen@ucl.ac.uk; coccia@lct.jussieu.fr; molivuc@bgnet.bgsu.edu; leonardo.guidoni@univaq.it

Abstract

The penta-2,4-dieniminium cation (PSB3) displays similar ground state and first excited state potential energy features as those of the retinal protonated Schiff base (RPSB) chromophore in rhodopsin. Recently, PSB3 has been used to benchmark several electronic structure methods, including highly correlated multireference wave function approaches, highlighting the necessity to accurately describe the electronic correlation in order to obtain reliable properties even along the ground state (thermal) isomerization paths. In this work, we apply two quantum Monte Carlo approaches, the variational Monte Carlo and the Lattice Regularized Diffusion Monte Carlo, to study the energetics and electronic properties of PSB3 along representative minimum energy paths and scans related to its thermal *cis-trans* isomerization. Quantum Monte Carlo is used in combination with the Jastrow Antisymmetrized Geminal Power ansatz, which guarantees an accurate and balanced description of the static electronic correlation, thanks to the multiconfigurational nature of the Antisymmetrized Geminal Power term, and of the dynamical correlation, due to the presence of the Jastrow factor explicitly depending on electron-electron distances. Along the two ground state isomerization minimum energy paths of PSB3, CASSCF calculations yield wave functions having either charge transfer or diradical character in proximity of the two transition state configurations. Here, we observe that at the quantum Monte Carlo level of theory, both the two transition states show only a leading charge transfer character, and the conical intersection is observed only if the path connecting the two transition states is extended beyond the diradical transition state, namely by increasing the bond-length-alternation (BLA). These findings are in good agreement with the results obtained by

*To whom correspondence should be addressed

†Dipartimento di Fisica, Sapienza-Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy

‡London Centre for Nanotechnology, University College London, London WC1E 6BT, United Kingdom

¶Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, Via Vetoio 2, 67010 L'Aquila, Italy

§Sorbonne Universités, UPMC Univ Paris 06 & CNRS, UMR 7616, Laboratoire de Chimie Théorique, F-75005 Paris, France

‖Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403, United States

⊥Dipartimento di Chimica, Universitá di Siena, via De Gasperi 2, I-53100 Siena, Italy

#Dipartimento di Scienze Fisiche e Chimiche, Università degli studi dell'Aquila, Via Vetoio 2, 67010 L'Aquila, Italy

MRCISD+Q calculations, and they demonstrate the importance of having an accurate description of the static and dynamical correlation when studying isomerization and transition states of conjugated systems.

1 Introduction

The retinal protonated Schiff Base (RPSB, represented in Fig. 1a) is the chromophore responsible for the photochemical properties of a vast family of biological photoreceptors referred to as retinal proteins which are, among other functions, involved in the mechanism of vision of dim light in vertebrates[1–3]. RPSB undergoes a very fast *cis-trans* isomerization in the protein (opsin) environment ($\sim$ 200 fs) with high quantum yield ($\sim$ 0.68) upon photon absorption[1]. This isomerization process has been deeply investigated by different spectroscopic techniques[4] and theoretical calculations[5,6]. In particular, femtosecond spectroscopy[7] and hybrid quantum mechanics / molecular mechanics (QM/MM) molecular dynamics calculations[7,8] highlight the essential role played by the surrounding protein environment and by the hydrogen-out-of-plane motion[9] in the isomerization mechanism, which involves decay via a S_1/S_0 conical intersection (CI)[10] reached by the selective torsion of the central ($C_{11}=C_{12}$) double bond.

The penta-2,4-dieniminium cation (PSB3, represented in Fig. 1a) is a small conjugated molecular system widely used as reduced computational model of the full RPSB, since PSB3 and RPSB exhibit similar ground and excited state features. First, both PSB3 and RPSB have equilibrium *cis* and *trans* isomers where the positive charge is localized on the nitrogen-containing side of the conjugated chain. Second, the transition from the ground state (S_0) to the first excited state (S_1) is characterized by a transfer of the positive charge from the nitrogen region towards the opposite end of the conjugated chain. Third, twisting the conjugated chain along one of the central double bonds leads to a CI structure between the S_0 and S_1 electronic states. Such a CI plays an important role in the photoisomerization mechanism of RPSB since it mediates population transfer from S_1 to S_0 along the isomerization coordinate.

Not only do PSB3 and RPSB have similar near-equilibrium and S_1 potential energy surface features, but it was recently found that they have similar ground state potential energy features in the vicinity of the CI as well. In a hybrid quantum mechanics/molecular mechanics (QM/MM) study of the thermal isomerization mechanism of bovine rhodopsin, it was found that thermal 11-*cis* to all-*trans* isomerization of RPSB may occur via one of two distinct saddle points that are in the vicinity of the S_0/S_1 CI[2]. Both saddle points (i.e., chemically, transition states) feature an almost orthogonally twisted $C_{11}=C_{12}$ double bond, similar to the CI, but very different bond length alternations (BLA). We define the BLA as the difference between the average bond length of formal single bonds (C1-C2 and C3-C4) and the average bond length of formal double bonds ($C_1=N$, $C_2=C_3$, and $C_4=C_5$), such that BLA is positive if there is no bond inversion and negative if there is. The situation is analogous in the case of PSB3. Indeed, PSB3 also features two transition states which are almost 90 degrees twisted along the central ($C_2=C_3$) double bond, but having different BLA patterns. In both PSB3 and RPSB, the two transition states not only have different BLA geometries, but also different electronic structures. One transition state (TSCT) is characterized by a transfer of the positive charge from the nitrogen-containing side of the molecule to the other end of the conjugated chain. The other transition state (TSDIR) retains the positive charge on the nitrogen side of the chain, and therefore has covalent/diradical character due to the homolytic breaking of the isomerizing double bond. In order to characterize the regions of the ground state potential energy surface driving the thermal isomerization, Gozem et al.[11] optimized the two transition states in PSB3 and used them to map three pathways at the CASSCF/6-31G* level of theory. The first two pathways are minimum energy paths (MEPs) leading away from each transition state and towards the two equilibrium structures, *cis*-PSB3 and *trans*-PSB3. These paths are called MEPCT (for the path passing through the charge transfer transition state TSCT) and MEPDIR (for the path passing through the covalent/diradical transition state TSDIR). The third path is an interpolation/extrapolation of coordinates between the two transition states. Since the transition states essentially have

different BLAs but otherwise are almost geometrically identical, this scan follows a BLA co-ordinate and is therefore called the BLA scan. This scan intercepts a S_1/S_0 CI structure. For a simplified representation of the three paths see Fig. 1b; further information are reported in Ref. 11.

The three aforementioned paths served useful as benchmarks to understand the importance of a balanced description of electronic correlation in correctly describing the topology of the potential energy surface in that region, as well as in computing the relative energies of TSCT and TSDIR. If the CASSCF(6,6)/6-31G* can be considered a good compromise between accuracy and computational effort for mapping (small) parts of the potential energy surface of complex molecular systems, using more sophisticated methodologies for single-point energy calculations on the CASSCF paths is seen to be a reliable strategy to explicitly include dynamic correlation in the wave function[11–13,15] allowing one to correct the original CASSCF findings such as CI location, energy of the π^2 and $\pi\pi^*$ configurations and relative stability of the TSCT and TSDIR.

Thanks to the tremendous progress in High Performance Computing (HPC) in recent years, Quantum Monte Carlo (QMC) methods represent a powerful alternative to other *ab-initio* and DFT approaches for the accurate description of systems where electronic correlation plays an essential role. QMC methods[16,17] have been widely applied to several problems of physical and chemical interest such as materials[18–22], molecular properties[23–34], reaction pathways[35,36] and geometry and excited states of biochromophores[37–41]. The good scalability with respect to the system size (N^d, with $3 < d < 4$ and N the number of electrons)[16,39] and the use of massively parallel algorithms make QMC methods particularly suitable for Petascale architectures. All these aspects justify the growing number of applications of QMC in problems of quantum chemistry and molecular physics.

The variational Monte Carlo (VMC) method[42] is the simplest QMC approach: thanks to a combined use of the Monte Carlo integration and the variational principle for the ground state, the many-body trial wave function Ψ_T can be efficiently optimized and used to ex-

Figure 1: Panel a: Lewis representation of the full chromophore, the 11-*cis* retinal protonated Schiff base (RPSB), covalently bound to a lysine of the opsin protein, and of the *cis*-, *trans*-penta-2,4-dieniminium cations (PSB3). Panel b: schematic representation of the thermal isomerization from the *cis* to the *trans* configuration of PSB3, following the two minimum energy paths (obtained in Ref. 11 from CASSCF calculations) undergoing respectively a charge transfer (CT) and a diradical (DIR) transition state. The former minimum energy path (MEPCT) is represented in red with a saddle point in the charge transfer transition state (TSCT), and the orange region indicates that the ground state potential energy surface is characterized by a charge transfer character. The latter minimum energy path (MEPDIR) is represented in blue with a saddle point in the diradical transition state (TSDIR), and the cyan region indicates that the ground state potential energy surface is characterized by a diradical character. The energy barriers corresponding to TSCT and TSDIR are comparable, and most of the computational studies agree that TSCT represents to lowest barrier.[11–13] The main difference between TSCT and TSDIR configurations lies in their bond length alternation (BLA), that is positive for the TSDIR and negative for the TSCT (see discussion in the text). A scan of the configurations connecting TSCT with TSDIR (called BLA scan) presents a conical intersection (CI) for a positive BLA, but the actual value varies with the adopted computational method.[11–14]
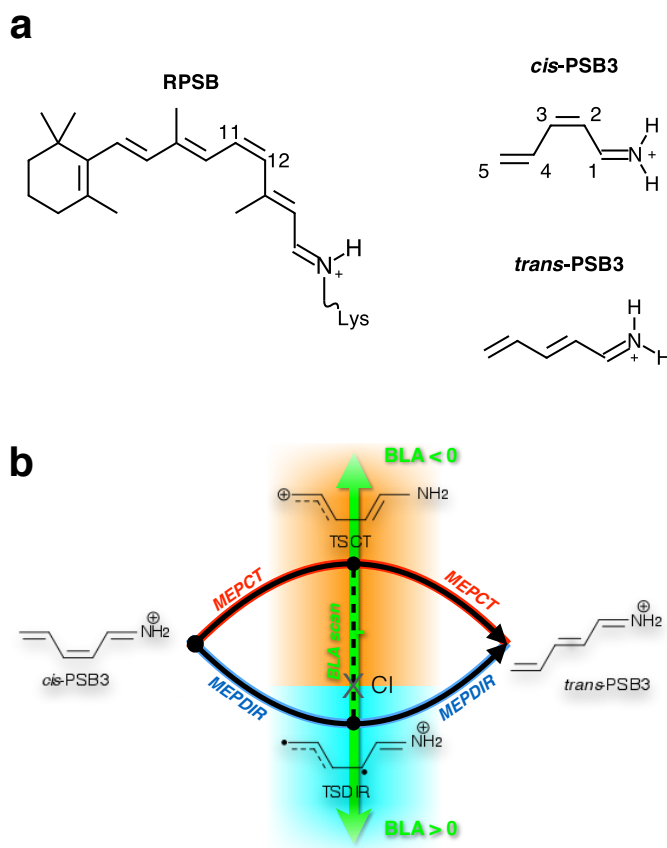
tract information other than just the energy of the molecular target. Higher accuracy in the determination of the system properties (energy, geometrical and electronic parameters, etc.) is usually achieved by the fixed node projection Monte Carlo methods, such as the diffusion Monte Carlo (DMC)[43,44] or the lattice regularized diffusion Monte Carlo[45,46] (LRDMC). The choice of the functional form of Ψ_T represents a crucial step determining the overall quality of QMC calculations, both in the variational and in the fixed node projection schemes. The Jastrow Antisymmetrized Geminal Power (JAGP)[47–52] has been seen to be very efficient in the investigation of chemical systems,[28–30,32,33,35,39–41] with accuracy comparable to that of high-level quantum chemistry methods. Its compactness, coupled to the use of efficient algorithms for the optimization of all parameters, including linear coefficients and exponents of the atomic basis set,[25,53,54] leads to a fast convergence of the variational results for electronic and geometrical properties with the size of the basis set[28,55,56], with a computational cost comparable to that of a simple wave function defined by the product of a Jastrow factor and a single Slater determinant. Although AGP is not size consistent, the presence of a flexible Jastrow factor makes JAGP size consistent for systems which splits in fragments with spin zero or $1/2$.[25,49,57] JAGP has already proven to give a good description of the static and dynamical correlation in several crucial benchmarks[32,50,51,58], like in the estimation of the torsional energy of the ethylene and of the singlet-triplet gap of methylene[32].

In the following, we compute the energy and electronic structure of PSB3 along the MEPCT, MEPDIR, and BLA paths reported previously[11]. As shown below, the combination of QMC and JAGP ansatz allows us to get a proper description of both static and dynamical correlation; for this reason, using the QMC and JAGP wave functions represents an optimal choice to study the intrinsic properties of the PSB3 model.

The paper is organized as follows: in Section 2 we report the main features of VMC and LRDMC schemes, and a detailed analysis of the JAGP ansatz and of its potential to correctly describe multiconfigurational systems, such as diradicals; computational details on our calculations are given in Section 3; the current results are shown in Section 4, pointing

out the importance of the electronic correlation properly introduced by the QMC/JAGP study for the characterization of the conical intersection; conclusions and comments on future perspectives are reported in the last Section.

2 Quantum Monte Carlo

2.1 Variational and Lattice Regularized Diffusion Monte Carlo

The accuracy of QMC approaches, both in the simplest VMC scheme and in the fixed-node projection schemes, are strictly related to the wave function ansatz. Typically, the electronic wave function Ψ_T in QMC[16,17,56] is defined by the product

$$\Psi_T(\bar{\mathbf{x}}; \bar{\mathbf{R}}) = \mathcal{D}(\bar{\mathbf{x}}; \bar{\mathbf{R}})\mathcal{J}(\bar{\mathbf{x}}; \bar{\mathbf{R}}), \tag{1}$$

where $\mathcal{D}$ is the antisymmetric function taking into account the fermionic nature of electrons and $\mathcal{J}$ is the Jastrow factor depending explicitly on the inter-particle (electrons and nuclei) distances; $\bar{\mathbf{x}}$ and $\bar{\mathbf{R}}$ represent the collective electronic ($\bar{\mathbf{x}}$ refers to space $\bar{\mathbf{r}}$ and spin $\bar{\sigma}$) and nuclear coordinates, respectively.

The Jastrow factor is a symmetric positive function of the electronic positions; therefore it does not change the nodal surface (determined by the antisymmetric term $\mathcal{D}$), but it introduces the dynamical correlation among electrons and satisfies the electron-electron and electron-nucleus cusp conditions[16,56,59].

In VMC, the parameters that define Ψ_T are optimized in order to minimize the electronic energy within the functional freedom of the ansatz. The VMC results can further be improved by using the fixed-node (FN) projection Monte Carlo techniques, which provide the lowest possible energy with the constraint that the wave function Φ_{FN} has the same nodal surface of an appropriately chosen guiding function Ψ_T[16,43], which is usually optimized using the VMC method. The fixed-node projection Monte Carlo method that we have adopted is

LRDMC[45,46], which is efficient for systems with a large number of electrons[46] and preserves the variational principle even when used in combination with nonlocal pseudopotentials[46]. Since the LRDMC calculations are much more demanding than the VMC calculations, in terms of computational time, they have been performed only for a few key structures.

2.2 The Jastrow Antisymmetrized Geminal Power

The trial wave function ansatz used in the QMC calculations presented in this paper is the Jastrow Antisymmetrized Geminal Power (JAGP),[47,48,56] that is the product

$$\Psi_T(\bar{\mathbf{x}}) \equiv \Psi_{JAGP}(\bar{\mathbf{x}}) = \Psi_{AGP}(\bar{\mathbf{x}}) \cdot \Psi_J(\bar{\mathbf{x}}) \tag{2}$$

of the Antisymmetrized Geminal Power (AGP) function Ψ_{AGP} and the Jastrow factor Ψ_J, where the dependance on the nuclear coordinates $\bar{\mathbf{R}}$ is here omitted.

For an unpolarized system (zero total spin S) of $N = 2N_p$ electrons and M atoms, the AGP function is defined as:

$$\Psi_{AGP}(\bar{\mathbf{x}}) = \hat{\mathcal{A}} \left[\prod_i^{N_p} G\left(\mathbf{x}_i; \mathbf{x}_{N_p+i}\right) \right], \tag{3}$$

where $\hat{A}$ is the antisymmetrization operator, and the geminal pairing function G is a product of a singlet function and a spatial wave function symmetric with respect to the particle exchange $\mathcal{G}$:

$$G(\mathbf{x}_i; \mathbf{x}_j) = \mathcal{G}\left(\mathbf{r}_i, \mathbf{r}_j\right) \frac{\alpha(i)\beta(j) - \beta(i)\alpha(j)}{\sqrt{2}}. \tag{4}$$

The spatial function $\mathcal{G}$ is a linear combination of products of atomic orbitals ϕ_μ:

$$\mathcal{G}\left(\mathbf{r}_i, \mathbf{r}_j\right) = \sum_\mu^L \sum_\nu^L g_{\mu\nu} \phi_\mu\left(\mathbf{r}_i\right) \phi_\nu\left(\mathbf{r}_j\right) \tag{5}$$

where the indexes μ and ν run over all the basis in all the atoms in the system, for a total of

L atomic orbitals (note that L is determined by the overall basis set size). The coefficients $g_{\mu\nu}$ have to be optimized in order to minimize the variational energy of the system (together with the other parameters in the wave function).

In our calculations we used this Jastrow factor

$$\Psi_J = \exp(U_{en} + U_{ee} + U_{een} + U_{eenn}),$$

that involves: the one-electron interaction term U_{en}, the homogeneous two-electron interaction term U_{ee}, and the inhomogeneous two-electron interaction terms U_{een} and U_{eenn} (representing respectively an electron-electron-nucleus function and an electron-electron-nucleus-nucleus function). They are defined as follows:

$$U_{en}(\bar{\mathbf{r}}) = \sum_a^M \sum_i^N \left[-Z_a \frac{1 - e^{-b_1 \sqrt[4]{2Z_a} r_{ia}}}{b_1 \sqrt[4]{2Z_a}} + \sum_\mu^{l_a} f_\mu^a \chi_\mu^a(\mathbf{r}_{ia}) \right] \tag{6}$$

$$U_{ee}(\bar{\mathbf{r}}) = \sum_{i<j}^N \left[\frac{1 - e^{-b_2 r_{ij}}}{2b_2} \right] \tag{7}$$

$$U_{een}(\bar{\mathbf{r}}) = \sum_a^M \sum_{i<j}^N \left[\sum_\mu^{l_a} \sum_\nu^{l_a} \bar{f}_{\mu,\nu}^a \chi_\mu^a(\mathbf{r}_{ia}) \chi_\nu^a(\mathbf{r}_{ja}) \right], \tag{8}$$

$$U_{eenn}(\bar{\mathbf{r}}) = \sum_{a \neq b}^M \sum_{i<j}^N \left[\sum_\mu^{l_a} \sum_\nu^{l_b} \tilde{f}_{\mu,\nu}^{a,b} \chi_\mu^a(\mathbf{r}_{ia}) \chi_\nu^b(\mathbf{r}_{jb}) \right], \tag{9}$$

where the vector $\mathbf{r}_{ia} = \mathbf{r}_i - \mathbf{R}_a$ is the difference between the position of the nucleus a and the electron i, r_{ia} is the corresponding distance, r_{ij} is the distance between electrons i and j, Z_a is the electronic charge of the nucleus a, which is described by l_a atomic orbitals χ_μ^a (with index $\mu = 1, \ldots, l_a$),[1] and b_1, b_2, $f_{\mu_a}^a$, $\bar{f}_{\mu,\nu}^a$, $\tilde{f}_{\mu,\mu}^{a,b}$ are variational parameters. The leading contribution for the description of electronic correlation is given by U_{ee}, but also the inhomogeneous two-electron interaction terms U_{een} and U_{eenn} are particularly important in

---

[1]Note that the atomic orbitals χ used in the Jastrow term are similar to the atomic orbitals ϕ used for the AGP, although they are not the same orbitals, and in general a reliable description of molecular systems requires a much smaller number of orbitals in the Jastrow than the number of orbitals used in the AGP, see for instance Ref. 56.

the JAGP ansatz, because they reduce the unphysical charge fluctuations included in the AGP function, as discussed in Refs. 25,49.

The pairing spatial function $\mathcal{G}$ in Eq. (5) is written in terms of the (localized) atomic orbitals ϕ_μ, offering an interesting correspondence between the AGP ansatz and the Resonating Valence Bond framework[60,61]. An equivalent way to write the pairing function $\mathcal{G}$ is obtained by using the molecular orbitals (MOs) ψ_k. The expansion of the pairing function in terms of MOs is obtained by performing a generalized (the atomic orbitals ϕ_μ are not necessarily orthonormal, so the overlap matrix $S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle \neq \delta_{\mu\nu}$) diagonalization of the coupling matrix $\mathbf{G}$, which is the $L \times L$ matrix of the $g_{\mu\nu}$ coefficients:

$$\mathbf{GSP} = \mathbf{P\Lambda} \tag{10}$$

$$\text{where } \Lambda = \text{diag}(\lambda_1, \dots, \lambda_L)$$

$$\text{and } |\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_L| \geq 0. \tag{11}$$

The resulting pairing function $\mathcal{G}$ is:

$$\mathcal{G}\left(\mathbf{r}_i, \mathbf{r}_j\right) = \sum_{k=1}^{L} \lambda_k \psi_k\left(\mathbf{r}_i\right) \psi_k\left(\mathbf{r}_j\right), \tag{12}$$

where the orthonormal single particle functions are written as

$$\psi_k(\mathbf{r}) = \sum_{\mu=1}^{L} P_{\mu k} \phi_\mu(\mathbf{r}) \tag{13}$$

with the $P_{\mu k}$ coefficients defining the eigenvectors $\mathbf{P}$.

It is interesting to investigate the connection between the expansion of the pairing function $\mathcal{G}$ in terms of MOs and the standard configuration interaction expansion of the wave function in multiconfigurational approaches.

By substitution of Eq. (12) in Eq. (3), and expanding the summation out of the antisym-

metrization operator, the following multi-determinant expansion is obtained for the AGP function:

$$\Psi_{AGP} = c_0 |\Psi_0\rangle + \sum_{i=1}^{N_p} \sum_{a=N_p+1}^{L} c_{ii}^{aa} |\Psi_{ii}^{aa}\rangle + \sum_{\substack{i,j=1 \\ i\neq j}}^{N_p} \sum_{\substack{a,b=N_p+1 \\ a\neq b}}^{L} c_{iijj}^{aabb} |\Psi_{iijj}^{aabb}\rangle + \ldots \qquad (14)$$

where the coefficients are given by:

$$c_0 = \prod_{i}^{N_p} \lambda_i \, ; \qquad c_{ii}^{aa} = c_0 \frac{\lambda_a}{\lambda_i} \, ; \qquad c_{iijj}^{aabb} = c_0 \frac{\lambda_a \lambda_b}{\lambda_i \lambda_j} \, ; \qquad \ldots, \qquad (15)$$

$|\Psi_0\rangle$ is the leading closed-shell Slater determinant:

$$|\Psi_0\rangle = \hat{\mathcal{A}} \left\{ \left[\prod_{i}^{N_p} \psi_i(\mathbf{r}_i)\alpha(i) \right] \left[\prod_{j}^{N_p} \psi_j(\mathbf{r}_{N_p+j})\beta(j) \right] \right\}, \qquad (16)$$

the determinant $|\Psi_{ii}^{aa}\rangle$ is equal to $|\Psi_0\rangle$, but with the virtual orbital ψ_a substituting the valence orbital ψ_i, etc. From the expression of the coefficients in Eq. (15) and the ordering of the eigenvalues λ_k in Eq (11), it follows that the leading contribution beyond the determinant $|\Psi_0\rangle$ is given by the determinant $|\Psi_{ii}^{aa}\rangle$ with $i = N_p$ and $a = N_p + 1$. The multideterminant expansion of Ψ_{AGP} in Eq. (14) allows us to directly compare the Ψ_{AGP} with wave functions from other quantum chemical frameworks. In Ψ_{AGP} all the odd excited determinants (single, triple, etc.) are excluded, whereas a subset of the even excitations (those with a multiple excitation to the same virtual orbital) are taken into account; only doubly occupied molecular orbitals are present. In other words, Ψ_{AGP} is contained in the seniority zero sector of the electronic full configuration interaction, and its expansion coefficients are determined by the ratios of the eigenvalues of the Λ matrix.

The seniority number Ω represents an alternative tool to classify singlet wave functions. Ω is defined as the number of unpaired electrons in the Slater determinant, e. g. the number of singly occupied molecular orbitals. In cases when the static correlation plays a major

role Ω-based selection of important Slater determinants in the expansion has been seen to be superior than the traditional one, based on the number of excitations with respect to the reference configuration.[62] Wave functions with $\Omega = 0$ for benchmark systems are accurate enough to recover most of the static correlation, but the FCI limit (including dynamic correlation) is achieved only when configurations from $\Omega = 2, 4, 6...$ sectors are explicitly included.[62] In the case of JAGP wave function, the combination between a $\Omega = 0$ determinantal term and a Jastrow factor allows us to estimate the correlation energy more accurately than $\Omega = 0$ CI wave functions. The set of MOs ψ_k is optimized within the JAGP framework, *i.e.* in presence of the Jastrow factor and of the multiconfigurational character of the wave function: MOs extracted from our optimization procedure represent therefore the optimal choice for the correlated description of the system under study. The way to move from the AGP MOs to the standard ones and vice versa is explained in the next Paragraph and in the Appendix.

2.3 AGP for diradicals

The multiconfigurational nature of the AGP function, clearly shown by Eq. 14, has been extensively discussed by Zen *et al.* in a recent paper on the use of AGP for diradicals[32]. Here we focus the attention on the application of the AGP ansatz to the $\pi\pi^*$ state of PSB3 (see the discussion in Section 4), starting from the simple but accurate model with two electrons in two orbitals, originally introduced by Salem and Rowland[63,64]. For a generic diradical system, two atomic orbitals ϕ_A and ϕ_B are centered on nuclei A and B. Such model is a representative scheme for molecules undergoing a double bond breaking, like the twisted ethylene or the twisted PSB3: in these cases, the two involved orbitals are of p type and positioned on two (central) carbon atoms.

The main goal of this discussion is to demonstrate that the AGP ansatz, a formally $\Omega = 0$ wave function, contains the $\pi\pi^*$ configuration, needed for a proper description of the ground

state surface of PSB3 along the paths.[2]

In case of two orbitals, the pairing function $\mathcal{G}$ term in the AGP formulation is explicitly given by

$$\mathcal{G}\left(\mathbf{r}_1, \mathbf{r}_2\right) = g_{AA}\phi_A(\mathbf{r}_1)\phi_A(\mathbf{r}_2) + g_{BB}\phi_B(\mathbf{r}_1)\phi_B(\mathbf{r}_2) + g_{AB}\phi_A(\mathbf{r}_1)\phi_B(\mathbf{r}_2) + g_{BA}\phi_B(\mathbf{r}_1)\phi_A(\mathbf{r}_2) \quad (17)$$

where $g_{\mu\nu}$ coefficients represent the coupling terms of the $\mathbf{G}$ matrix in the expansion of the AGP spatial factor: g_{AA} and g_{BB} are referred to the ionic terms $\phi_A(\mathbf{r}_1)\phi_A(\mathbf{r}_2)$ and $\phi_B(\mathbf{r}_1)\phi_B(\mathbf{r}_2)$ in which the two electrons are localized on the same atom, whereas the elements g_{AB} and g_{BA} are related to the covalent terms $\phi_A(\mathbf{r}_1)\phi_B(\mathbf{r}_2)$ and $\phi_B(\mathbf{r}_1)\phi_A(\mathbf{r}_2)$.

In terms of molecular orbitals, $\mathcal{G}\left(\mathbf{r}_1, \mathbf{r}_2\right)$ becomes ($L = 2$)

$$\mathcal{G}\left(\mathbf{r}_1, \mathbf{r}_2\right) = \lambda_1 \psi_1(\mathbf{r}_1)\psi_1(\mathbf{r}_2) + \lambda_2 \psi_2\left(\mathbf{r}_1\right) \psi_2(\mathbf{r}_2). \quad (18)$$

On the other hand, in the standard delocalized picture[63,64] two ψ_+ and ψ_- molecular orbitals are defined:

$$\begin{aligned} \psi_+ &= \frac{\phi_A + \phi_B}{\sqrt{2}} \\ \psi_- &= \frac{\phi_A - \phi_B}{\sqrt{2}} \end{aligned} \quad (19)$$

assuming zero overlap between ϕ_A and ϕ_B (the assumption is easily verified for two orthogonal p orbitals).

The diradical $\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_+\psi_-$ wave function ($\psi_+\psi_-$ is a short notation for $\psi_+(\mathbf{r}_1)\psi_-(\mathbf{r}_2) + \psi_-(\mathbf{r}_1)\psi_+(\mathbf{r}_2)$) is written in the atomic basis as the following

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{\phi_A(\mathbf{r}_1)\phi_A(\mathbf{r}_2) - \phi_B(\mathbf{r}_1)\phi_B(\mathbf{r}_2)}{\sqrt{2}} \quad (20)$$

---

[2]The π and π^* orbitals have to be intended here as π-molecular orbitals residing on highly twisted configurations.

Since the AGP formally contains the terms introduced in Eq. 20, the comprehension of a consistent way to link together Eqs. 17, 18 and 20 is a mandatory task. Some questions are therefore arising: i) which relation occurs between ψ_1, ψ_2, ψ_+ and ψ_- ; ii) which relation exists between the atomic basis, ϕ_A and ϕ_B, and the molecular orbitals, ψ_1 and ψ_2 obtained by the diagonalization of the $\mathcal{G}$ matrix (Eq. 18); iii) finally, if a diradical ground state $\psi_+\psi_-$ can be properly described by the AGP ansatz.

An unitary matrix $\hat{\mathcal{U}}$ transforms the ψ_+ and ψ_- molecular orbitals, by rotating them with a certain angle θ, into $\tilde{\psi}_+$ and $\tilde{\psi}_-$

$$
\begin{pmatrix} \tilde{\psi}_+ \\ \tilde{\psi}_- \end{pmatrix} = \hat{\mathcal{U}} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \tag{21}
$$

If $\theta = \pi/4$, the normalized orbitals become

$$
\tilde{\psi}_+ = \frac{1}{\sqrt{2}} (\psi_+ + \psi_-) \tag{22}
$$

$$
\tilde{\psi}_- = \frac{1}{\sqrt{2}} (\psi_+ - \psi_-) . \tag{23}
$$

By defining $\psi_1 \equiv \tilde{\psi}_+$ and $\psi_2 \equiv \tilde{\psi}_-$, which become $\psi_1 = \phi_A$ and $\psi_2 = \phi_B$ by taking into account the relations in Eqs. 19, and by substitution in Eq. 18, we obtain

$$
\mathcal{G}(\mathbf{r}_1, \mathbf{r}_2) = \lambda_1 \phi_A(\mathbf{r}_1)\phi_A(\mathbf{r}_2) + \lambda_2 \phi_B(\mathbf{r}_1)\phi_B(\mathbf{r}_2), \tag{24}
$$

which is equivalent to Eq. 20 for $\lambda_1 = -\lambda_2 = 1/\sqrt{2}$. We also observe that Eq. 17 is equivalent to Eqs. 20 for $g_{AB} = g_{BA} = 0$ and $g_{AA} = -g_{BB} = 1/\sqrt{2}$.

Summarizing, it is always possible to transform a $\psi_+\psi_-$ configuration into a combination of $\tilde{\psi}_+^2$ and $\tilde{\psi}_-^2$ by applying an opportunely chosen proper unitary transformation. In this simple model for diradicals, the AGP molecular orbitals are a linear combination of the orbitals deriving from the traditional picture of electronic delocalization. The detailed derivation of

Figure 2: Comparison between CASSCF frontier (π and π^*) and AGP frontier (ψ_1 and ψ_2, see Eq. 18) orbitals; corresponding combinations, given the unitary transformation $\hat{\mathcal{U}}$ (Eq. 21), are also shown.
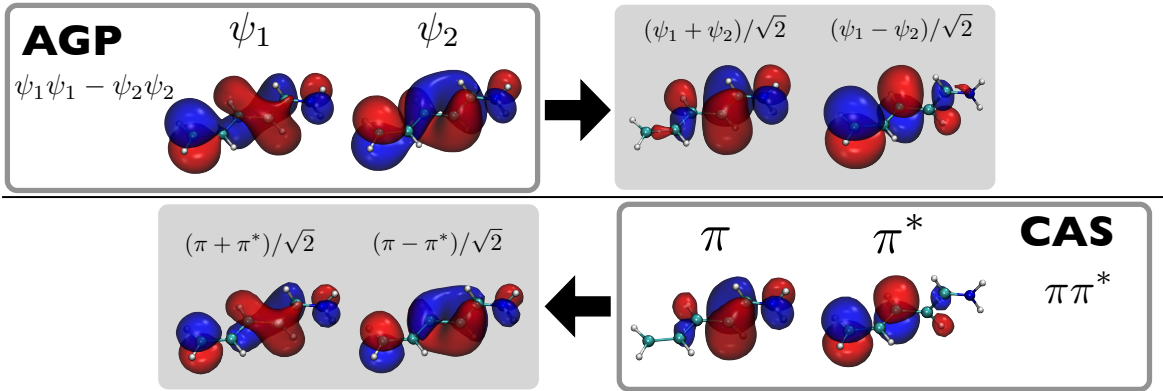
the mapping procedure is reported in the Appendix A, where the generalization from the (2,2) to the (2,n) active space (where n is the number of orbitals) is shown and the proof of the capability of AGP to represent such active space is reported. The expansion in Eq. 17 shows that the AGP ansatz contains all the terms reported by the picture in terms of delocalized molecular orbitals and localized atomic orbitals. The $g_{\mu\nu}$ coefficients are variational parameters optimized by the stochastic methods mentioned before and for this reason the AGP optimization is a fundamental step to select the right wave function for the ground state of interest.

As a conclusion of this discussion, the graphical representation of CASSCF and AGP molecular orbitals for highly twisted diradical $\pi\pi^*$ state of a certain structure of PSB3 (the last structure of the BLA path, see the discussion below), together with the combinations explained in this paragraph, is reported in Fig. 2: the AGP frontier orbitals (ψ_1 and ψ_2, Eq. 18) are identical to the normalized sum and difference of the corresponding CASSCF orbitals (π and π^* orbitals), according to the rotation given by the unitary transformation $\hat{\mathcal{U}}$.

3 Computational details

The QMC calculations reported in this paper have been obtained using the *TurboRVB* package developed by S. Sorella and coworkers[65], that includes a complete suite of variational and diffusion Monte Carlo codes for wave function and geometry optimization of molecules and solids. The scalar-relativistic energy consistent pseudopotentials (**ECP**) of Burkatzki *et al.*[66] have been used in order to describe the two core electrons of the carbon and nitrogen atoms. In detail, the basis sets we have used for the AGP part are: (10s,9p,2d,1f) contracted in {8} hybrid orbitals for the carbon atom, (8s,9p,2d,1f) contracted in {8} hybrid orbitals for the nitrogen atom, (6s,5p,1d) contracted in {1} hybrid orbitals for the hydrogen atom. As basis sets for the atomic orbitals included in the inhomogeneous terms of the Jastrow factor, namely in U_{en}, U_{een} and U_{eenn} reported in Eqs. 6, 8 and 9, we used an uncontracted basis for the U_{en} and U_{een} term, and a contracted with hybrid orbitals basis for U_{eenn}. This allowed us to have an accurate basis set for the Jastrow factor while keeping the number of parameters of the wave function reasonably small. In more details, in U_{en} and U_{een} we used a (4s,2p,1d) basis set for the carbon or nitrogen atoms, and a (3s,2p) for hydrogen atom, whereas in U_{eenn} the orbitals are contracted in {2} hybrid orbitals for the oxygen, nitrogen or hydrogen atoms.

The parameters of the wave function Ψ_T, including the values of the exponents of the atomic orbitals, have been optimized by using the already validated and stable optimization schemes discussed in Ref. 56. In particular, the optimization that we have followed for the singlet π^2 and triplet states of PSB3 considered here in the different structures, starts from an initial configuration where the AGP matrix is diagonal, the exponents are initialized to values taken from standard Dunning's basis sets (where too small and too large values are eliminated because they are not not necessary due to the presence of our Jastrow factor, see discussion in Ref. 56) and all the Jastrow parameters are set to zero, with the exception of $b_1 = b_2 = 1$. Next, the optimization procedure follows the protocol: *(i)* optimization of the AGP, namely of the matrix elements and the contraction coefficients of the basis set,

with fixed exponents and Jastrow parameters $b_1 = b_2 = 1$; *(ii)* optimization of the AGP and relaxation of the values of the exponents of the AGP basis set and of the b_1 and b_2 parameters; *(iii)* optimization of the Jastrow terms, keeping the AGP parameters fixed; *(iv)* optimization of the overall JAGP, keeping fixed the exponents in the basis set, both for the AGP and the Jastrow; *(v)* optimization of all the parameters, including the exponents of the basis set, with increasing statistical accuracy. For the single diradical state ($\pi\pi^*$ configuration) of PSB3 in the proximity of the conical intersection, we have used a slightly different procedure to avoid the possibility to be trapped in a local minimum. Therefore, we forced the wave function to be in the correct electronic configurations by taking the triplet JAGP optimized wave function, and obtained from that the corresponding singlet diradical configuration. This wave function has been used as the starting point of an optimization that started form the step *(iii)* of the previously stated optimization schedule. We have verified *a posteriori* for every nuclear structure where we have calculated both the π^2 and the $\pi\pi^*$ singlet configurations, that the overlap $\langle \Psi_{JAGP}^{\pi\pi^*} | \Psi_{JAGP}^{\pi^2} \rangle$ between the two JAGP wave functions (calculated using the correlated sampling techniques) is almost zero, thus the two wave functions actually correspond to different electronic states.

It is important to note that QMC approaches use stochastic methods both to evaluate an observable, and to optimize the variational parameters of the wave function. The stochastic uncertainty due to the former point is easy to calculate and it has been reported in figures and tables of the present work. The latter point is instead much more difficult to evaluate. We have carefully tested the reliability of the optimization schemes used in this work, and indeed the profiles reported in the following figures, although not perfectly smooth, are pretty regular, both for energy and charge transfer values. Moreover, the most interesting configurations (*cis*, *trans*, and structures close to TSCT and TSDIR) have been optimized with some extra-effort, thus the results reported in the tables are fully reliable.

In this work we also report several results computed at the level of the fixed-node projection Monte Carlo scheme, that has been realized by performing LRDMC calculations with

mesh size $a = 0.3$ a.u. Although we have not performed, for computational reasons, the continuous extrapolation of the lattice mesh size $a \to 0$, we know from previous works[32,56] and preliminary calculations that the bias given by the finite mesh size $a = 0.3$ a.u. is almost negligible in the evaluations of the considered energy differences.

4   Results and Discussion

In this section we present the energetic and electronic features of the MEPCT, MEPDIR and BLA paths computed using VMC and LRDMC with the JAGP ansatz. A high-level treatment of electron correlation is crucial for the correct description of the energy surface of the ground state isomerization of PSB3. As discussed in Ref. 11, dynamic electron correlation modifies the mapped CASSCF potential energy surface in two ways: the TSCT transition state is found lower in energy than the TSDIR, at variance with the CASSCF findings, and the CASSCF CI is seen shifted to larger BLA values when dynamic correlation is included in the calculations. The energies along the three paths have been computed using a number of electronic structure methods, namely: multiconfigurational approaches[11], multi-reference perturbation theory[11], DFT schemes[13,67], EOM-CC[12] and SORCI[15]. All the most accurate methods qualitatively predict similar changes in the potential energy surface with respect to CASSCF. In particular, in the case of MRCISD+Q (the most accurate method tested previously), TSCT becomes more stable than TSDIR by 4.7 kcal/mol (compared to CASSCF where it is less stable by 1.2 kcal/mol), and the CI gets shifted to a BLA value of ~ 0.03 Å (compared to ~ 0.00 Å for CASSCF).

Before starting, we validated our computational protocol by looking at the electronic properties of the *cis* isomer of PSB3 such as the dipole moment μ and the charge-transfer of the ground state S_0, defined as the partial charge on the allyl moiety $H_2C=CH-CH=$ (the net charge of the system is +1). The charge transfer character at *cis*-PSB3 is 0.313 at the CASSCF/6-31G* level of theory and 0.355 at the MRCISD+Q level of theory, as derived

from Mulliken population analyses[11] (see Tab. 1). Since we cannot define Mulliken charges in our QMC framework, we compute the charge transfer by finding the portion of the electronic density in the region of the allyl moiety up to the plane perpendicular to the $C_2=C_3$ bond and cutting it in the middle. This method of obtaining the charge-transfer character is tested on densities extracted from DFT and wave function methods and is shown to produce very similar charge transfer character as Mulliken charges.

The single-reference nature of the S_0 state is highlighted by the fact that charge-transfer values computed with standard DFT (with PBE and B3LYP functionals) and HF are in good agreement with the MRCISD+Q result. The effect of applying pseudopotentials on the carbon and nitrogen atoms on the charge transfer is found to be negligible when comparing all-electron (AE) and ECP results obtained using similar basis sets, as shown in Tab. 1. A further evidence of the reliability of our approach is given by the performance of several variants of the single-determinant (SD) wave function ($J_{1\text{-body}}$SD, where only the 1-body term for the Jastrow is used; JSD-proj, the SD wave function is projected out from the full AGP; JSD-opt, the SD wave function is optimized after projection) in the VMC framework (e. g., taking into account only the first term of the AGP expansion in Eq. 14) that, using ECP, are in full agreement with the more accurate VMC/JAGP (0.358, the best variational result) and the MRCISD+Q. LRDMC/JAGP only slightly corrects (0.352) the VMC/JAGP result for the charge-transfer. The same conclusions are easily extended to the dipole moment.

4.1 MEPCT and MEPDIR paths

As first, we consider the energy difference between the *cis* and *trans* isomers of PSB3. VMC/AGP and LRDMC/AGP values (Tab. 2), -2.9(2) and -3.0(2) kcal/mol respectively, and the VMC/JSD value (-2.9(2) kcal/mol) are fully consistent with the reference MRCISD+Q (-3.1 kcal/mol)[11,12], XMCQDPT2 (-2.8 kcal/mol)[11] and EOM-CCSD (-3.0 kcal/mol)[12]. The

Table 1: Singlet ground-state energy (in Hartree, H), total dipole μ (in Debye, D) and charge-transfer value for the *cis* isomer of PSB3, evaluated with several computational methods, as defined in the first column. $J_{1\text{-body}}SD$, JSD-proj and JSD-opt are defined in the text, and the basis sets for the QMC calculations are defined in Section 3. In the Core column, AE stands for all-electron calculation, ECP for energy-conserving pseudo-potential calculation. The reported numbers in the last column represent the net charge on the allyl moiety. VMC and LRDMC errors on μ and charge-transfer are of the order of 10^{-3}.

Method	Ref.	Core	Energy [H]	μ [D]	Charge-transfer
CASSCF(6,6)//6-31G*	11	AE	—	—	0.313
MRCISD+Q//6-31G*	11	AE	—	—	0.355
PBE//cc-pVTZ	this work	AE	—	3.784	0.375
PBE//VTZ-ANO	this work	ECP	—	3.758	0.369
B3LYP//cc-pVTZ	this work	AE	—	3.718	0.380
B3LYP//VTZ-ANO	this work	ECP	—	3.639	0.376
HF//cc-pVTZ	this work	AE	—	3.441	0.405
HF//VTZ-ANO	this work	ECP	-41.7258	3.472	0.396
VMC/SD^a	this work	ECP	-41.7048(9)	3.625	0.365
VMC/J$_{1\text{-body}}$SD^b	this work	ECP	-41.7130(7)	3.633	0.365
VMC/JSD-proj	this work	ECP	-42.8361(2)	3.895	0.363
VMC/JSD-opt	this work	ECP	-42.8373(2)	3.900	0.360
VMC/JAGP	this work	ECP	-42.8490(2)	3.983	0.356
LRDMC/JAGP	this work	ECP	-42.9160(3)	4.066	0.352

^{a} Wave function optimization by DFT/LDA; $E_{\text{DFT/LDA}} = -42.6663848$ H.
^{b} Wave function optimization by DFT/LDA; $E_{\text{DFT/LDA}} = -42.6769218$ H.

Table 2: Energy differences ΔE (in kcal/mol) between the singlet ground state of *cis* PSB3 isomer and: the singlet ground state *trans* isomer, the TSCT and TSDIR obtained by Gozem et al.[11] from CASSCF-based calculations. The energy difference ΔE_{S-T} between the first singlet and the first triplet electronic states of the *cis* isomer is also reported. The reported QMC results are compared with MRCISD+Q, XMCQDPT2 and EOM-CCSD calculations.

Method	Ref.	ΔE *trans*	ΔE TSCT	ΔE TSDIR	ΔE_{S-T} *cis*
VMC/JSD	This work	-2.9(2)	44.7(2)	51.1(2)	62.9(2)
VMC/JAGP	This work	-2.9(2)	45.2(2)	51.7(2)	66.2(2)
LRDMC/JAGP	This work	-3.0(2)	45.5(2)	51.4(2)	63.9(2)
MRCISD+Q	11,12	-3.1	48.7	54.9	—
XMCQDPT2	11	-2.8	46.9	50.5	—
EOM-CCSD	12	-3.0	46.6	52.5	—

negative value indicates that *trans*-PSB3 is more stable than *cis*-PSB3, since throughout this work the *cis*-PSB3 energy is taken as the reference.

At VMC/JAGP level, the TSCT (45.5(2) kcal/mol) is lower in energy than the TSDIR (51.7 kcal/mol), making the CT path energetically favored, similarly to what as reported by the aforementioned correlated approaches[11,12]. LRDMC and VMC findings are equal within the stochastic error to the VMC/JAGP values, evidence that the trial wave function Ψ_T is fully optimized. The singlet-triplet gap for the *cis*-PSB3 is also reported, with a difference of 2.3(3) kcal/mol between VMC and LRDMC using the complete JAGP.

The VMC/JAGP energy profile of the MEPCT path (Fig. 3a) is characterized by a shallow plateau around the transition state structure, at variance with the shape of the MEPDIR path (Fig. 3c); moreover, expensive LRDMC calculations do not alter the picture. Panels b) and d) of Fig. 3 show the ratio between λ_{LUMO} and λ_{HOMO}, according to the AGP expansion given in Eq. 14; as already discussed by some of us in the case of the application of the AGP ansatz on the diradical twisted ethylene C_2H_4[32], very small values of this ratio correspond to a single-reference wave function, with the lowest molecular orbitals doubly occupied (the ratio is exactly zero in the limit of a pure single Slater determinant), whereas large values of the ratio indicate two (near)-equivalent configurations. Following the analysis in Ref. 32, the absolute value of $\lambda_{\text{LUMO}}/\lambda_{\text{HOMO}}$ approaching to unity means that two configurations are contributing with the same weight to the electronic structure. λ_{HOMO} and λ_{LUMO} correspond

to λ_1 and λ_2 in the simple (2,2) model in Eq. 18. The wave function along the MEPCT path (Fig. 3b) is dominated by a single configuration, with the $\lambda_{\text{LUMO}}/\lambda_{\text{HOMO}}$ oscillating around the value for the *cis* isomer (-0.0635), as expected by the previous investigations on the MEPCT path of PSB3. Similar arguments can be found in Ref. 67. The same behaviour has been found along the MEPDIR path (Fig. 3b), with no appreciable contribution given by higher-energy configurations: this result unequivocally shows that the introduction by QMC of balanced dynamic correlation strongly alters the CASSCF description of the electronic structure of PSB3 along the MEPDIR path, similarly to what found in MRCISD+Q calculations.

The AGP wave function can be formally expanded into a linear combination of Slater determinants (Eq. 14), so the single-electron molecular orbitals are obtained by the diagonalization of the geminal coupling matrix (Eq. 10) and are defined within a correlated framework. The AGP spans the seniority number $\Omega = 0$ sector in Hilbert space, with double occupation for each orbital. The terms "closed-shell" and "open-shell" are widely used to indicate systems without and with unpaired electrons, respectively, implying that the molecular orbitals come from single-reference approaches, like Hartree-Fock or DFT. In the case of highly twisted configurations of PSB3, for instance, the π^2 state, involving charge transfer with respect to the equilibrium ground state, has closed-shell character, while the diradical $\pi\pi^*$ state has open-shell character: for $\pi\pi^*$ static correlation plays an important role, and a multi-configurational approach must be used. Such definitions strictly depend on the choice of the reference for the molecular orbitals, as explained in Section 2.3. Even though the AGP wave function is formally characterized by only doubly occupied orbitals, its application is not limited so far to the study of closed-shell systems since the molecular orbitals involved in the AGP expansion are the results of a variational optimization and they can be qualitatively different from the Hartree-Fock ones, as explicitly shown in Section 2.3. Data show that the MEPDIR path does not have a diradical character anymore; the VMC/JAGP description of MEPCT and MEPDIR paths produces a ground state of closed-shell π^2 char-

acter, similarly to the MRCISD+Q[13].

The results above are again reflected in the charge transfer profiles along both MEPCT and MEPDIR paths shown in Fig. 4. Indeed, it is clear that in both MEPCT and MEPDIR, the wave function gains charge transfer character, with TSCT and TSDIR both having ca. 65% of the positive charge on the allyl moiety. The picture presented at the VMC/JAGP and LRDMC/JAGP levels is again at variance with what is found at the CASSCF level of theory. With CASSCF, MEPDIR passes through a transition state, TSDIR, that has diradical character and has almost no charge transfer at all, with all the charge localized on the nitrogen-containing moiety. This difference is due to the change in CI position on the energy surface after introducing a balanced description of electronic correlation in the calculations. At the CASSCF level of theory, the CI is peaked and lies in between the TSCT and TSDIR transition states, causing them to have wave functions with different character. At correlated levels of theory, the CI is shifted to a larger BLA value than that corresponding to TSDIR. As a result, both TSCT and TSDIR lie on the same side of the CI (which is now intermediate/sloped) and therefore both have the same wavefunction character (charge transfer). This change in local topology causes only one of the transition states to remain as a saddle point, TSCT, while TSDIR is no longer a transition state on the S_0 potential energy surface. The charge transfer character in both MEPCT and MEPDIR decreases as PSB3 moves towards the *cis* or *trans* isomers of PSB3, converging to the values corresponding to the two minima.

Furthermore, the MEPCT maxima (i.e. the TSCT) region is flatter for the two QMC levels than for MRCISD+Q level; this may be due to a better treatment of the electron dynamic correlation in this region.

4.2 BLA path

In the CASSCF(6,6)/6-31G* landscape the CI is located between the TSCT and TSDIR. Consistently with the other approaches, PSB3 assumes a closed-shell, charge transfer char-

Figure 3: Left column: a) thermal isomerization energy profile (with respect to the *cis* PSB3) and b) ratio between λ_{LUMO} and λ_{HOMO} (see Eq. 14) along the MEPCT reaction coordinate. Right column: c) thermal isomerization energy profile and d) ratio between λ_{LUMO} and λ_{HOMO} along the MEPDIR reaction coordinate. SD stands for Single Determinant, e. g. a single configuration (with ratio $\lambda_{\text{LUMO}}/\lambda_{\text{HOMO}}$ exactly equal to zero). Error bars are within the symbols. For both paths CASSCF and MRCISD+Q energy profiles are taken from Refs. 11,12.
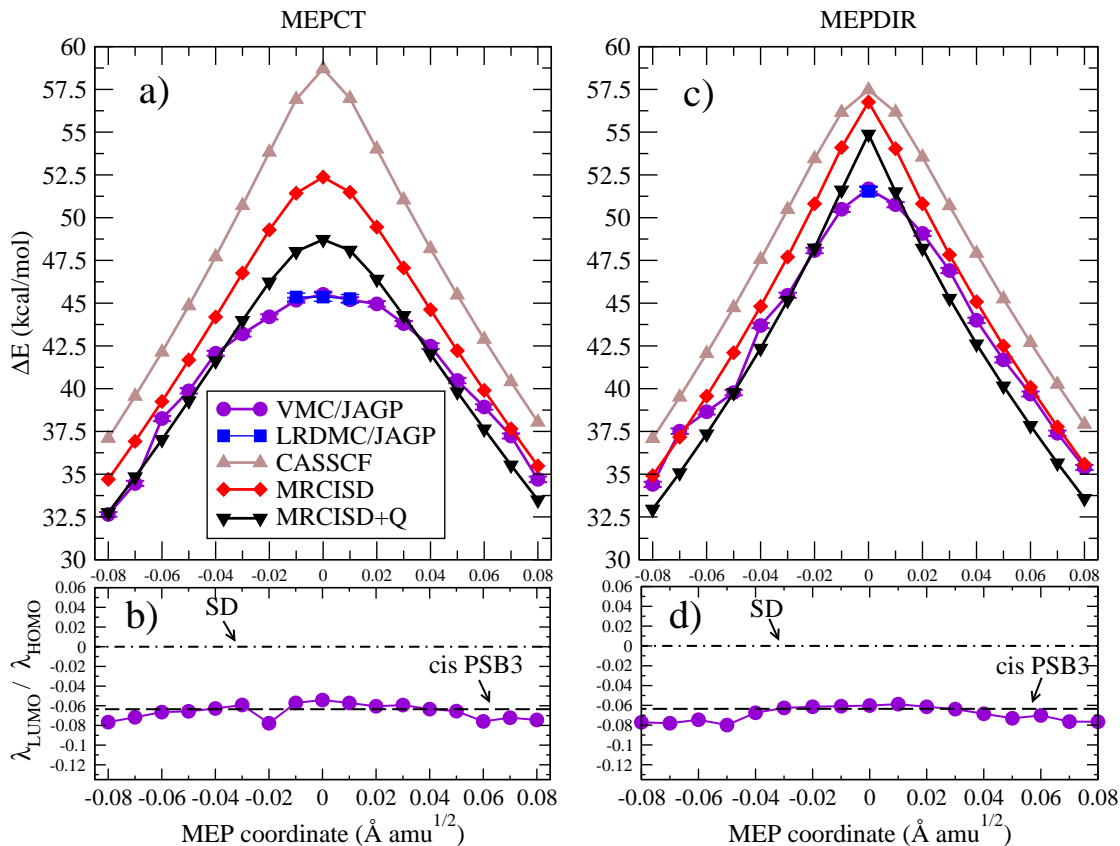
Figure 4: Charge-transfer along the MEPCT and MEPDIR paths. CASSCF and MRCISD charge-transfer profiles are taken from Refs. 11,12.
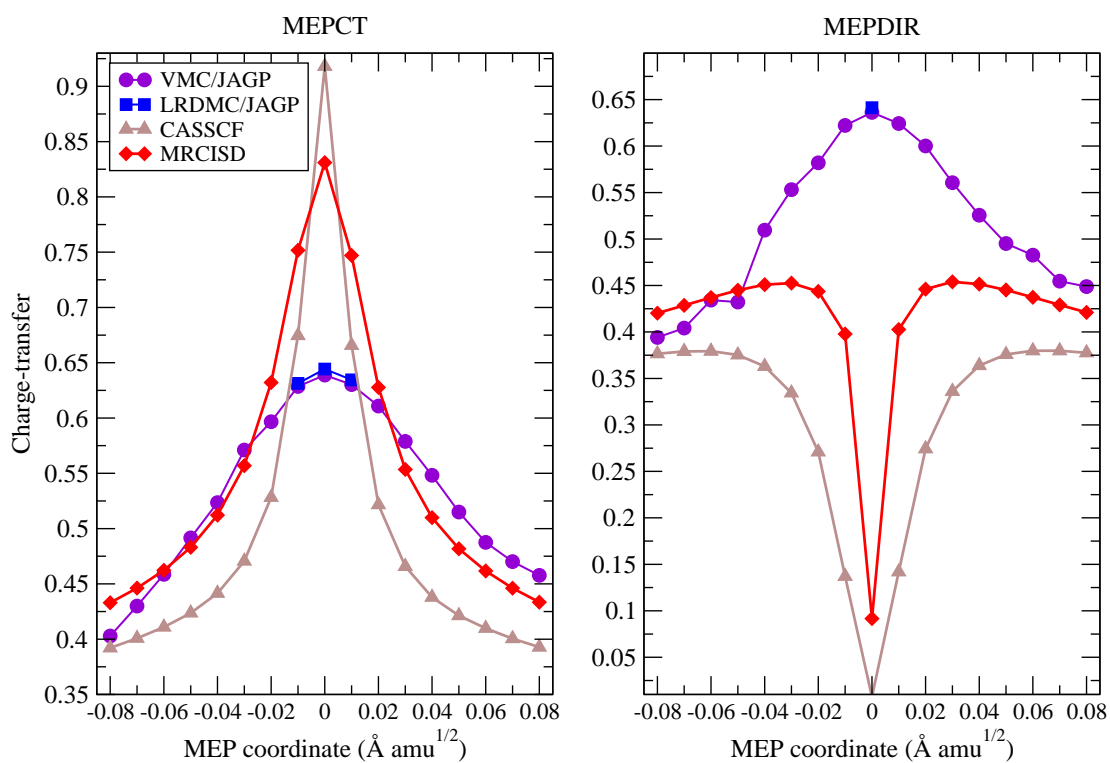
acter (π^2) at smaller BLA values than the CI, while its wave function becomes covalent and diradical ($\pi\pi^*$) at larger BLA values (upper panel of Fig. 5). As anticipated above, the crossing between the two states does not produce a peaked CI, like in the CASSCF case, but instead a sloped and intermediate CI. Even though we are not able to identify the exact position of the CI, VMC/JAGP calculations clearly show that CI moves towards large values of BLA ($\sim$0.075 Å), beyond the TSDIR. The present result differs from the collection of data obtained by other correlated approaches, locating the CI around 0.03-0.04 Å[11–13,15], with the exception of the QD-NEVPT2/CAS(6,6) analysis (extrapolated value at 0.05 Å)[11]. When using the LRDMC/JAGP approach, the CI position ($\sim$0.06 Å) comes closer to MR-CISD+Q data. We observe that the charge transfer (π^2) curves obtained using VMC/JAGP and LRDMC/JAGP are actually more stable – with respect to *cis*-PSB3 energy – than that obtained with MRCISD+Q, consistently with the lower energy TSCT of VMC/JAGP and LRDMC/JAGP. Meanwhile, the diradical ($\pi\pi^*$) curve from VMC/JAGP or LRDMC/JAGP are less stable that those from MRCISD+Q. This is what causes the CI from VMC/JAGP and LRDMC/JAGP to shift to higher BLA than in MRCISD+Q. Moreover, it is important to point out that relaxing the structures and the minimum energy paths of PSB3 at the QMC correlated level may produce slightly different topologies for the BLA and MEP scans close to the CI.

In the lower panel of Fig. 5 the VMC/JAGP points corresponding to the $\pi\pi^*$ state are characterized by a high ratio (in absolute value) between λ_{HOMO} and λ_{LUMO}, close to 1. As already mentioned, this is due to the multi-configurational nature of the diradical state, where the frontier orbitals are near-degenerate and singly occupied. From the AGP analysis, near-degeneracy for S_0 is therefore found in the large-BLA portion of the path. This reinforces the results of the MEPCT and MEPDIR paths above, since it is clear here that TSCT and TSDIR both lie on the same side of the CI after introduction of the dynamic electron correlation. Also, one can see here that while TSCT remains a saddle point on the S_0 potential energy surface, TSDIR actually becomes a minimum on the S_1 surface and is

no longer a transition state as in CASSCF. The π^2 state, also extending in the diradical portion of space previously defined by the CASSCF study, is clearly single-reference, and it dominates the wave function of PSB3 up to a BLA of 0.05 Å.

The reliability of the present results is further confirmed looking at the triplet energy profile along the BLA path (blue triangles in the upper panel of Fig. 5, filled and open symbols for VMC and LRDMC, respectively). Since the comparison of the π^2 and $\pi\pi^*$ energies along the BLA scan is a fundamental step in order to give an accurate representation of the ground state potential energy surface of PSB3 surrounding the CI, the triplet energy profile can be considered a lower bound for the $\pi\pi^*$ energy. For diradicals, triplet should be the ground-state spin multiplicity, according to the molecular version of Hund's rule[63,64]: this is the case, for instance, of the orthogonally twisted ethylene molecule[32] where, due to the homolytic cleavage of the double bond, the wave function is dominated by two configurations with the same weight and the two involved p atomic orbitals have zero overlap. The structures of the BLA path with diradical character are characterized by a torsion of about 90 degrees around the formal central double bond of PSB3, similar to the prototypical example given by C_2H_4 system: the central double bond is broken and the two p orbitals are mutually (almost) perpendicular. This manifest similarity between the twisted diradical PSB3 and the prototypical C_2H_4 system yields to reasonably expect that the ground state of the twisted diradical PSB3 is also a triplet[3], as we observe *a posteriori* with VMC/JAGP and LRDMC/JAGP calculations.

As one can see from the upper panel of Fig. 5, the triplet energy is higher than the π^2 singlet energy for BLA values smaller than 0.055 Å for VMC and 0.05 Å for LRDMC, and consequently the same certainly occurs for the $\pi\pi^*$ state. A similar gap of $\sim$ 10-12 kcal/mol is also found at CASSCF(6,6)/6-31G* level in the diradical portion of the BLA path.

These simple arguments, combined with the fact that the convergence of the wave function

---

[3]However, it is not true in general for all diradicals that the triplet state has a a lower energy that the singlet state: for example, disjointed and non-Kekule molecules, like the tetramethyleneethane[31,68], have a very small singlet-triplet gap with the singlet lower in energy in some specific geometries, while the ground state of oligacenes larger than hexacene[69] is also a singlet.

optimization for a triplet state is easier to achieve in the VMC/AGP framework, allow us to be extremely confident with the robustness of the present results for the location of CI along the BLA path.

The ground state wave function is dominated by the π^2 state before the CI (i.e., at lower

Figure 5: Energy profile along the BLA path (with respect to the *cis* PSB3) and the related $\lambda_{\text{LUMO}}/\lambda_{\text{HOMO}}$ ratio for the π^2 and $\pi\pi^*$ configurations, calculated by VMC and LRDMC methods on the variationally optimized JAGP wave functions. The triplet energy profile is also reported. The MRCISD+Q, MRCISD and CASSCF profiles are shown, for comparison, and the conical intersections (CI) obtained with the different approaches are marked in the plot.
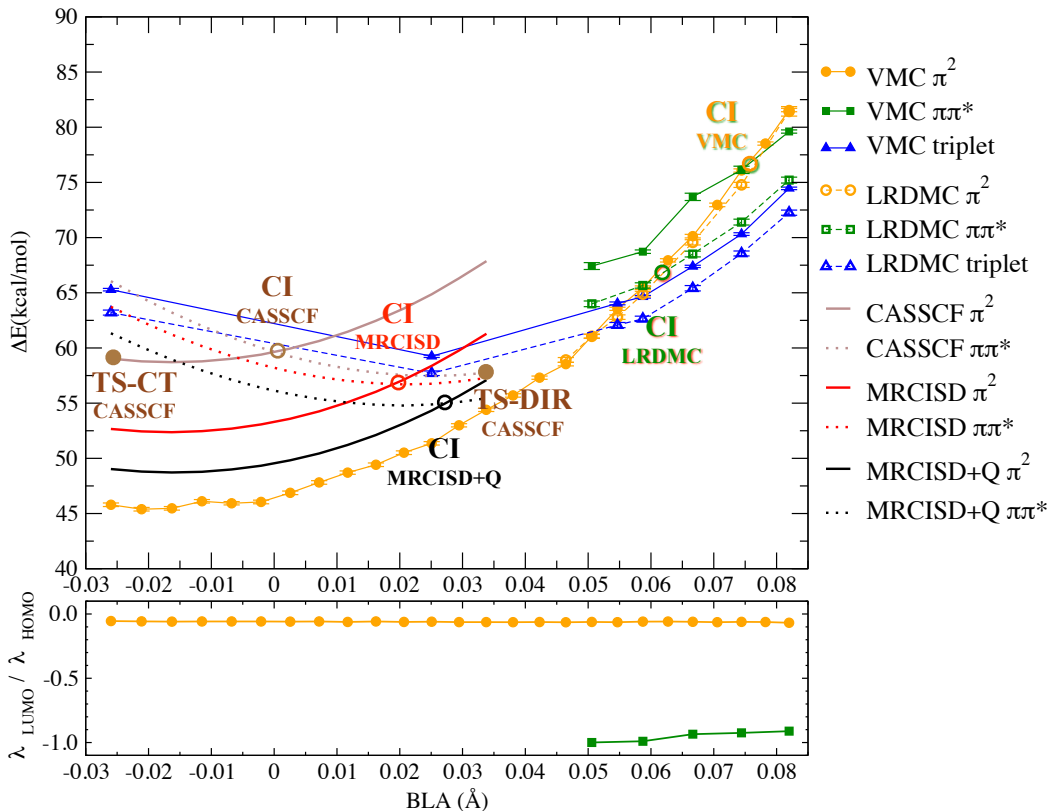
BLA values), while it assumes $\pi\pi^*$ character after the crossing; it is interesting to understand how the charge distribution changes in the two configurations. Fig. 6 reports, as an example, the charge density difference between the π^2 and $\pi\pi^*$ states for VMC/JAGP and LRDMC/JAGP CI points along the BLA path, more precisely, for the closest structures to CI found at VMC and LRDMC level. The shape of the isosurfaces obviously resembles the p

atomic orbitals involved in the π electronic structure of the twisted PSB3. The most evident changes in the charged distribution are observed for the two central carbon atoms, where the double bond has been broken because of the torsion: the yellow part of the plot (online color version) corresponds to an accumulation of electronic charge in the π^2 state with respect to the $\pi\pi^*$, whereas the green isosurfaces indicate the opposite situation.

Such analysis is consistent with the charge transfer nature of π^2 configuration and with the covalent and diradical character of $\pi\pi^*$. We know from Fig. 7 that the charge-transfer of π^2 along the path oscillates between 0.63 and 0.72, and that the charge-transfer of $\pi\pi^*$ is much smaller (~ 0.10). A large (small) value of charge-transfer means an excess (reduction) of electronic charge on the protonated imine heteroallyl moiety, according to the plot reported in Fig. 7 for VMC and LRDMC calculations, with the charge-transfer of the triplet coinciding with the $\pi\pi^*$ values.

5 Conclusions

VMC and LRDMC methods have been used to compute electronic and energetic properties of PSB3, using the JAGP ansatz.

VMC/JAGP and LRDMC/JAGP calculations on three different CASSCF paths reveal the fundamental role played by a balanced description of the dynamic correlation for the correct representation of the ground state energy surface in the proximity of the conical intersection. The VMC and LRDMC calculations significantly alter the CASSCF landscape, inverting the relative stability of the MEPCT and MEPDIR paths, similarly to what obtained by other correlated approaches.[11–13,15] However the region surrounding the TSCT appears more flat when computed at the QMC levels.

On both MEPCT and MEPDIR paths, the PSB3 wave function assumes a charge transfer and single-reference character: the MEPDIR path is therefore not diradical anymore
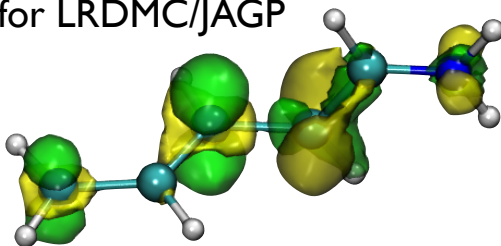
Figure 6: Graphical representation of the charge density difference between the π^2 and the $\pi\pi^*$ configurations for VMC/JAGP and LRMDC/JAGP CI points: the yellow (green) isosurface indicates an excess (reduction) of 0.01 in terms of electronic charge in the π^2 state with respect to $\pi\pi^*$.

CI for VMC/JAGP
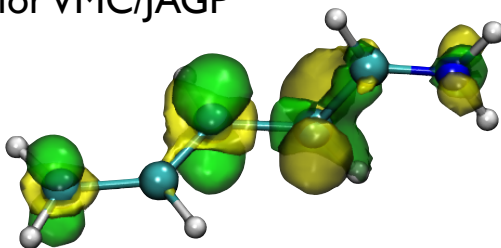
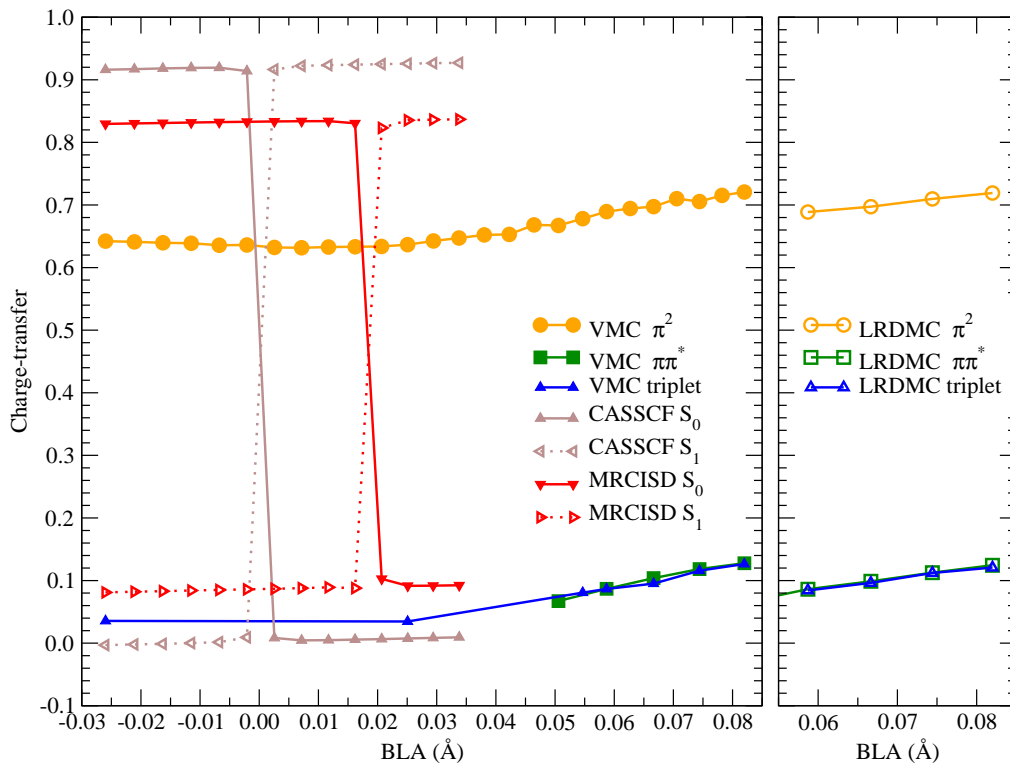CI for LRDMC/JAGP

$\rho_{\pi^2} < \rho_{\pi\pi^*}$ $\rho_{\pi^2} > \rho_{\pi\pi^*}$

Figure 7: Charge-transfer along the BLA path for the π^2, $\pi\pi^*$ and triplet configurations, obtained by VMC and LRDMC approaches (stochastic errors are smaller than point size). The MRCISD and CASSCF profiles are shown, for comparison.

when analyzed at correlated level. This finding is immediately confirmed by the fact the CI is pushed towards values of BLA larger than that corresponding to the TSDIR ($\sim$0.075 Å with VMC/JAGP and $\sim$0.06 Å with LRDMC/JAGP, while the reference calculations give a CI located at around 0.03-0.04 Å).

Furthermore, in the present study the multi-configurational nature of AGP has been explained in detail, following the work reported in Ref. 32, for the description of the diradical $\pi\pi^*$ configuration of PSB3. Even though the AGP is defined in the seniority number $\Omega = 0$ subsector of the Hilbert space (e.g., the set of molecular orbitals is doubly occupied), an unitary transformation of the AGP-optimized molecular orbitals allows one to always map the $\pi\pi^*$ state of highly twisted PSB3 geometries, with two unpaired electrons, into the AGP $\Omega = 0$ subsector.

The role of the dynamic electronic correlation has been found to be essential in order to get a reliable description of the ground state of PSB3 around the CI. Therefore, mechanistic or dynamics studies using methods which do not incorporate these effects need to be performed and interpreted critically. The recent improvements in forces calculations using QMC approaches[70] suggest that there will be soon the possibility to compute MEPCT and MEPDIR minimum energy paths based on the VMC/JAGP method, yielding to a more consistent comparison with the energies and geometries obtained from CASSCF, in order to further clarify the main features of PSB3.

A   Appendix: comparison between a CASSCF and a AGP wave function ansätze

In this appendix we will show the relationships between the singlet AGP wave function and the complete active space CAS ansätze. In particular, we will demonstrate the equivalence of a CAS(2,n) of 2 electrons in n molecular orbitals and a geminal function for a singlet system of 2 electrons. Later we will consider the case of an AGP function with two nearly

degenerate states.

A CAS(2,n) wave function for a singlet state is a linear combination of Slater determinants, where every Slater determinant is identified by the occupation numbers of some reference orthonormal molecular orbitals ψ_i, $i = 1, \ldots, n$ and the spin state of the two electrons, and the coefficients of the expansion are variationally optimized. If we indicate the possible combinations of 2 electrons in n orbitals with the second quantization notation:

$$
\begin{aligned}
|2, 0, \ldots, 0\rangle &= \hat{a}_{1,\uparrow}^\dagger \hat{a}_{1,\downarrow}^\dagger |0\rangle \\
|\alpha, \beta, \ldots, 0\rangle &= \hat{a}_{1,\uparrow}^\dagger \hat{a}_{2,\downarrow}^\dagger |0\rangle \\
|\beta, \alpha, \ldots, 0\rangle &= \hat{a}_{2,\uparrow}^\dagger \hat{a}_{1,\downarrow}^\dagger |0\rangle \\
|0, 2, \ldots, 0\rangle &= \hat{a}_{2,\uparrow}^\dagger \hat{a}_{2,\downarrow}^\dagger |0\rangle
\end{aligned}
\tag{25}
$$

$$\ldots$$

where $|0\rangle$ is the empty space, and the operator $\hat{a}_{i,\uparrow}^\dagger$ ($\hat{a}_{i,\downarrow}^\dagger$) creates an electron of spin up (α) or down (β) in the orbitals ψ_i, and satisfies the canonical anticommutation relations. Using this notation, a CAS(2,n) wave function can be simply written as:

$$
|\text{CAS}(2, n)\rangle = \sum_{i=1}^n c_{ii} \hat{a}_{i,\uparrow}^\dagger \hat{a}_{i,\downarrow}^\dagger |0\rangle + \sum_{i<j} c_{ij} \left(\hat{a}_{i,\uparrow}^\dagger \hat{a}_{j,\downarrow}^\dagger + \hat{a}_{j,\uparrow}^\dagger \hat{a}_{i,\downarrow}^\dagger \right) |0\rangle
\tag{26}
$$

where the coefficients c_{ii} and c_{ij} are variationally optimized. In Eq. 26 we have implicitly used the fact that the coefficients for the configurations $\hat{a}_{i,\uparrow}^\dagger \hat{a}_{j,\downarrow}^\dagger |0\rangle$ and $\hat{a}_{j,\uparrow}^\dagger \hat{a}_{i,\downarrow}^\dagger |0\rangle$ have to coincide in order to have a singlet wave function.

We will prove now that the CAS(2,n) wave function can be generally rewritten in the following form:

$$
|\text{CAS}(2, n)\rangle = \sum_{i=1}^n \lambda_i \hat{b}_{i,\uparrow}^\dagger \hat{b}_{i,\downarrow}^\dagger |0\rangle
\tag{27}
$$

where $\hat{b}_{i,\uparrow}^\dagger$ and $\hat{b}_{i,\downarrow}^\dagger$ are respectively the creation operators of a spin up and down electron in a orbital $\tilde{\psi}_i$ obtained from a unitary transformation of the orbitals $\psi_1, \ldots, \psi_n$. In other

words, we will prove that the second sum in the right hand size (RHS) of Eq. 26 can be drop by transforming the reference molecular orbitals with an opportunely chosen unitary transformation. Since the RHS of Eq. 27 is actually the expression of a geminal in terms of molecular orbitals (or of its natural orbitals), this will prove that CAS(2,n) is equivalent to the geminal function that we are using in the AGP framework.

Proof: Eq. 26 can be written in the following way:

$$\hat{\mathbf{a}}_\uparrow^\dagger \cdot \mathbf{C}\hat{\mathbf{a}}_\downarrow^\dagger|0\rangle \tag{28}$$

where $\hat{\mathbf{a}}_\uparrow^\dagger$ ($\hat{\mathbf{a}}_\downarrow^\dagger$) is a vector whose elements are the creation operators $\hat{a}_{i,\uparrow}^\dagger$ ($\hat{a}_{i,\downarrow}^\dagger$), $\mathbf{C}$ is the $n \times n$ symmetric matrix whose elements are c_{ij}, and "·" is a scalar product. The matrix $\mathbf{C}$ is diagonalized by a unitary matrix $\mathbf{U}$, such that $\mathbf{C} = \mathbf{U}^\dagger \Lambda \mathbf{U}$, being Λ a diagonal matrix with diagonal elements $\lambda_1, \ldots, \lambda_n$ and the symbol † indicates the conjugate transpose (in the case of having the matrix $\mathbf{C}$ real, as in this work, we have that $\mathbf{U}$ is also real, thus it is actually an orthogonal matrix). Eq. 28 is easily rewritten as:

$$\hat{\mathbf{b}}_\uparrow^\dagger \cdot \Lambda\hat{\mathbf{b}}_\downarrow^\dagger|0\rangle \tag{29}$$

with $\hat{\mathbf{b}}_\uparrow^\dagger \equiv \mathbf{U}\hat{\mathbf{a}}_\uparrow^\dagger$ and $\hat{\mathbf{b}}_\downarrow^\dagger \equiv \mathbf{U}\hat{\mathbf{a}}_\downarrow^\dagger$, and the elements of the vector $\hat{\mathbf{b}}_\uparrow^\dagger$ ($\hat{\mathbf{b}}_\downarrow^\dagger$) are thus the creation operators $\hat{b}_{i,\uparrow}^\dagger$ ($\hat{b}_{i,\downarrow}^\dagger$) of an up (down) electron in the orbital $\tilde{\psi}_i = \sum_{j=1}^n U_{ij}\psi_j$. The unitary transformed orbitals $\tilde{\psi}_i$ are orthonormal because $\mathbf{U}$ is unitary. It is straightforward to show that (29) is equal to the RHS of Eq. 27, and this concludes the proof. ∎

If we define the operator $\hat{G}$ that creates a singlet pair of electrons:

$$\hat{G} = \sum_{i=1}^n \lambda_i \hat{b}_{i,\uparrow}^\dagger \hat{b}_{i,\downarrow}^\dagger, \tag{30}$$

then, by comparing with Eq. 27, we have that $|\text{CAS}(2,n)\rangle = \hat{G}|0\rangle$. On the other hand, an AGP wave function $|AGP\rangle$ for a singlet system with N electrons, so $N/2$ singlet pairs, can

be written as:

$$|AGP\rangle = \hat{G}^{N/2}|0\rangle. \tag{31}$$

Without los of generality, we can assume that the orbitals $\tilde{\psi}_i$ such that their corresponding λ_i in Eq. 27 are ordered in decreasing order for their absolute value:

$$|\lambda_1| \le |\lambda_2| \le \ldots \le |\lambda_n|. \tag{32}$$

By substitution of (30) into the RHS of Eq. 31, and dropping out the terms that are zero due to the anticommutation relations of the $b^\dagger$ operators, and an irrelevant multiplicative coefficient $c = \prod_{i=1}^{N/2} \lambda_i$, we have that the AGP wave function can be expanded in the following way (in agreement with Ref. 32):

$$
\begin{aligned}
|AGP\rangle \quad \propto \quad & \prod_{i=1}^{N/2} b_{i,\uparrow}^\dagger b_{i,\downarrow}^\dagger |0\rangle + \\
& + \sum_{1 \le j \le N/2} \sum_{N/2 < p \le n} \frac{\lambda_p}{\lambda_j} \left(\prod_{\substack{i=1 \\ i \ne j}}^{N/2} b_{i,\uparrow}^\dagger b_{i,\downarrow}^\dagger \right) b_{p,\uparrow}^\dagger b_{p,\downarrow}^\dagger |0\rangle + \\
& + \sum_{1 \le j < k \le N/2} \sum_{N/2 < p < q \le n} \frac{\lambda_p \lambda_q}{\lambda_j \lambda_k} \left(\prod_{\substack{i=1 \\ k \ne i \ne j}}^{N/2} b_{i,\uparrow}^\dagger b_{i,\downarrow}^\dagger \right) b_{p,\uparrow}^\dagger b_{p,\downarrow}^\dagger b_{q,\uparrow}^\dagger b_{q,\downarrow}^\dagger |0\rangle + \\
& + \ldots
\end{aligned}
\tag{33}
$$

For the specific case of diradical molecules, AGP can provide a reliable description of the system by having $n = N/2 + 1$, and $|\lambda_{N/2+1}| \sim |\lambda_{N/2}| \ll |\lambda_{N/2-1}|$. Moreover, for a singlet system of N electrons, the AGP ansatz is comparable to the complete active space $CAS(2, n_o)$ of 2 electrons in n_o orbitals, if in Eq. 33 we have: $n = N/2 + n_o$, and $|\lambda_{N/2}| \ll |\lambda_{N/2-1}|$. In this way we have for $j \le N/2$ and $p > N/2$ that $|\lambda_p/\lambda_j| \le \epsilon$ (where $\epsilon \equiv |\lambda_{N/2}/\lambda_{N/2-1}|$ is small) and the absolute value of the coefficients of the third term in the RHS of Eq. 33 are $\left|\frac{\lambda_p \lambda_q}{\lambda_j \lambda_k}\right| \le \epsilon^2$, thus negligible. Analogous considerations lead to show that at higher orders of

expansion the coefficients are of order ϵ^m, with m growing with the order of the expansion, therefore for small ϵ the only non-negligible terms in the AGP expansion are the first and second one in the RHS of Eq. 33.

Acknowledgement

References

(1) Palczewski, K. *Annuv. Rev. Biochem.* **2006**, *75*, 743–767.

(2) Gozem, S.; Schapiro, I.; Ferré, N.; Olivucci, M. *Science* **2012**, *337*, 1225–1228.

(3) Ernst, O. P.; Lodowski, D. T.; Elstner, M.; Hegemann, P.; Brown, L. S.; Kandori, H. *Chem. Rev.* **2014**, *114*, 126–163.

(4) Kukura, P. *Science* **2005**, *310*, 1006–1009.

(5) Röhrig, U. F.; Guidoni, L.; Laio, A.; Frank, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2004**, *126*, 15328–15329.

(6) Röhrig, U. F.; Guidoni, L.; Rothlisberger, U. *Biochemistry* **2002**, *41*, 10799–10809.

(7) Polli, D.; Altoé, P.; Weingart, O.; Spillane, K. M.; Manzoni, C.; Brida, D.; Tomasello, G.; Orlandi, G.; Kukura, P.; Mathiers, R. A.; Garavelli, M.; Cerullo, M. *Nature* **2010**, *467*, 440–442.

(8) Frutos, L. M.; Andruniow, T.; Santoro, F.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7764–7769.

(9) Schapiro, I.; Ryazantsev, M. N.; Frutos, L. M.; Ferré, N.; Olivucci, M. *J. Am. Chem. Soc.* **2011**, *133*, 3354–3364.

(10) Andruniow, T.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17908–17913.

(11) Gozem, S.; Huntress, M.; Schapiro, I.; Lindh, R.; Granovsky, A. A.; Angeli, C.; Olivucci, M. *J. Chem. Theory Comput.* **2012**, *8*, 4069–4080.

(12) Gozem, S.; Krylov, A. I.; Olivucci, M. *J. Chem. Theory Comput.* **2013**, *9*, 284–292.

(13) Xu, X.; Gozem, S.; Olivucci, M.; Truhlar, D. G. *J. Phys. Chem. Lett.* **2013**, *4*, 253–258.

(14) Gozem, S.; Melaccio, F.; Valentini, A.; Filatov, M.; Huix-Rotllant, M.; Ferré, N.; Frutos, L. M.; Angeli, C.; Krylov, A. I.; Granovsky, A. A.; Lindh, R.; Olivucci, M. *J. Chem. Theory Comput.* **2013**, *10*.

(15) Schapiro, I.; Neese, F. *Comput. Theoret. Chem.* **2014**, in press.

(16) Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. *Rev. Mod. Phys.* **2001**, *73*, 33–83.

(17) Austin, B. M.; Zubarev, D. Y.; Lester, W. A. J. *Chem. Rev.* **2012**, *112*, 263–288.

(18) Spanu, L.; Sorella, S.; Galli, G. *Phys. Rev. Lett.* **2009**, *103*, 196401.

(19) Maezono, R.; Drummond, N. D.; Ma, A.; Needs, R. J. *Phys. Rev. B* **2010**, *82*, 184108.

(20) Kolorenc, J.; Mitas, L. *Rep. Prog. Phys.* **2011**, *74*, 026502.

(21) Mazzola, G.; Zen, A.; Sorella, S. *J. Chem. Phys.* **2012**, *137*, 134112.

(22) Mazzola, G.; Yunoki, S.; Sorella, S. *Nat. Comm.* **2014**, *5*, 1–6.

(23) Caffarel, M.; Rerat, M.; Pouchan, C. *Phys. Rev. A* **1993**, *47*, 3704–3717.

(24) Schautz, F.; Filippi, C. *J. Chem. Phys.* **2004**, *120*, 10931–10941.

(25) Sorella, S.; Casula, M.; Rocca, D. *J. Chem. Phys.* **2007**, *127*, 014105.

(26) Sterpone, F.; Spanu, L.; Ferraro, L.; Sorella, S.; Guidoni, L. *J. Chem. Theory Comput.* **2008**, *4*, 1428–1434.

(27) Zimmerman, P. M.; Toulouse, J.; Zhang, Z.; Musgrave, C. B.; Umrigar, C. J. *J. Chem. Phys.* **2009**, *131*, 124103.

(28) Barborini, M.; Sorella, S.; Guidoni, L. *J. Chem. Theory Comput.* **2012**, *8*, 1260–1269.

(29) Zen, A.; Zhelyazov, D.; Guidoni, L. *J. Chem. Theory Comput.* **2012**, *8*, 4204–4215.

(30) Coccia, E.; Chernomor, O.; Barborini, M.; Sorella, S.; Guidoni, L. *J. Chem. Theory Comput.* **2012**, *8*, 1952–1962.

(31) Pozun, Z. D.; Su, X.; Jordan, K. D. *J. Am. Chem. Soc.* **2013**, *135*, 13862–13869.

(32) Zen, A.; Coccia, E.; Luo, Y.; Sorella, S.; Guidoni, L. *J. Chem. Theory Comput.* **2014**,

(33) Zen, A.; Trout, B. L.; Guidoni, L. *J. Chem. Phys.* **2014**, *141*, 014305.

(34) Luo, Y.; Zen, A.; Sorella, S. *J. Chem. Phys.* **2014**, *141*, 194112.

(35) Barborini, M.; Guidoni, L. *J. Chem. Phys.* **2012**, *137*, 224309.

(36) Saccani, S.; Filippi, C.; Moroni, S. *J. Chem. Phys.* **2013**, *138*, 84109.

(37) Valsson, O.; Filippi, C. *J. Chem. Theory Comput.* **2010**, *6*, 1275–1292.

(38) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. *J. Chem. Theory Comput.* **2012**, *8*, 112–124.

(39) Coccia, E.; Guidoni, L. *J. Comput. Chem.* **2012**, *33*, 2332–2339.

(40) Coccia, E.; Varsano, D.; Guidoni, L. *J. Chem. Theory Comput.* **2013**, *9*, 8–12.

(41) Coccia, E.; Varsano, D.; Guidoni, L. *J. Chem. Theory Comput.* **2014**, *10*, 501–506.

(42) Bressanini, D.; Reynolds, P. J. *Advances in Chemical Physics, Monte Carlo Methods in Chemical Physics* **1998**, *105*, 5345–5350.

(43) Reynolds, P. J.; Ceperley, D. M.; Alder, B. J.; Lester, W. A. *J. Chem. Phys.* **1982**, *77*, 5593.

(44) Hammond, B. L.; Lester, W. A., Jr.; Reynolds, P. J. *Monte Carlo Methods in Ab-Initio Quantum Chemistry*; World Scientific, 1994.

(45) Casula, M.; Filippi, C.; Sorella, S. *Phys. Rev. Lett.* **2005**, *95*, 100201.

(46) Casula, M.; Moroni, S.; Sorella, S.; Filippi, C. *J. Chem. Phys.* **2010**, *132*, 154113.

(47) Casula, M.; Sorella, S. *J. Chem. Phys.* **2003**, *119*, 6500–6511.

(48) Casula, M.; Attaccalite, C.; Sorella, S. *J. Chem. Phys.* **2004**, *121*, 7110–7126.

(49) Neuscamman, E. *Phys. Rev. Lett.* **2012**, *109*, 203001–5.

(50) Neuscamman, E. *J. Chem. Phys.* **2013**, *139*, 181101.

(51) Neuscamman, E. *J. Chem. Phys.* **2013**, *139*, 194105.

(52) Stella, L.; Attaccalite, C.; Sorella, S.; Rubio, A. *Phys. Rev. B* **2011**, *84*, 245117.

(53) Sorella, S. *Phys. Rev. B* **2005**, *71*, 241103.

(54) Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Hennig, R. G. *Phys. Rev. Lett.* **2007**, *98*, 110201.

(55) Petruzielo, F. R.; Toulouse, J.; Umrigar, C. J. *J. Chem. Phys.* **2011**, *134*, 064104.

(56) Zen, A.; Luo, Y.; Sorella, S.; Guidoni, L. *J. Chem. Theory Comput.* **2013**, *9*, 4332–4350.

(57) Marchi, M.; Azadi, S.; Casula, M.; Sorella, S. *J. Chem. Phys.* **2009**, *131*, 154116.

(58) Stella, L.; Attaccalite, C.; Sorella, S.; Rubio, A. *Phys. Rev. B* **2011**, *84*, 245117–6.

(59) Drummond, N. D.; Towler, M. D.; Needs, R. J. *Phys. Rev. B* **2004**, *70*, 235119.

(60) Casula, M.; Yunoki, S.; Attaccalite, C.; Sorella, S. *Comput. Phys. Commun.* **2005**, *169*, 386–393.

(61) Sorella, S.; Zen, A. *The New Resonating Valence Bond Method for Ab-Initio Electronic Simulations*; Springer International Publishing: Cham, 2014; pp 377–392.

(62) Bytautas, L.; Henderson, T. M.; Jiménez-Hoyos, C. A.; Ellis, J. K.; Scuseria, G. E. *J. Chem. Phys.* **2011**, *135*, 044119.

(63) Salem, L.; Rowland, C. *Angew. Chem. Int. Ed.* **1972**, *11*, 92–111.

(64) Abe, M. *Chem. Rev.* **2013**, *113*, 7011–7088.

(65) Sorella, S. *TurboRVB* Quantum Monte Carlo package (accessed date May 2013). `http://people.sissa.it/~sorella/web/index.html`.

(66) Burkatzki, M.; Filippi, C.; Dolg, M. *J. Chem. Phys.* **2007**, *126*, 234105.

(67) Huix-Rotllant, M.; Filatov, M.; Gozem, S.; Schapiro, I.; Olivucci, M.; Ferré, N. *J. Chem. Theory Comput.* **2013**, *9*, 3917–3932.

(68) Chattopadhyay, S.; Chaudhuri, R. K.; Mahapatra, U. S. *Chem. Phys.* **2011**, *12*, 2791–2797.

(69) Bendakov, M.; Duong, H. M.; Starkey, K.; Houk, K.; Carter, E. A.; Wudl, F. *J. Am. Chem. Soc.* **2004**, *126*, 7416–7417.

(70) Sorella, S.; Capriotti, L. *J. Chem. Phys.* **2010**, *133*, 234111.