Reference Excitation Energies of Increasingly Large Molecules: A QMC Study of Cyanine Dyes

Alice Cuzzocrea, Saverio Moroni,* Anthony Scemama,* and Claudia Filippi*

ABSTRACT: We revisit here the lowest vertical excitations of cyanine dyes using quantum Monte Carlo and leverage recent developments to systematically improve on previous results. In particular, we employ a protocol for the construction of compact and accurate multideterminant Jastrow-Slater wave functions for multiple states, which we have recently validated on the excited-state properties of several small prototypical molecules. Here, we obtain quantum Monte Carlo excitation energies in excellent agreement with high-level coupled cluster for all the cyanines where the coupled cluster method is applicable. Furthermore, we push our protocol to longer chains, demonstrating that quantum Monte Carlo is a viable methodology to establish reference data at system sizes which are hard to reach with other high-end approaches of similar accuracy. Finally, we determine which ingredients are key to an accurate treatment of these challenging systems and rationalize why a description of the excitation based on only active $\pi$ orbitals lacks the desired accuracy for the shorter chains.

1. INTRODUCTION

Cyanine dyes are a family of charged $\pi$-conjugated molecules which are employed in very diverse applications ranging from dye-synthesized solar cells to the labeling of biomolecules.\textsuperscript{1-3} Their characteristic structure consists of a chain of an odd number of carbons with two amine groups at the ends. While their photophysical properties are strongly regulated by the length of the carbon chain, the lowest bright state of the cyanines always maintains a $\pi \rightarrow \pi^*$ character and can be predominantly described as a HOMO to LUMO (HL) transition. Despite the apparent simplicity of this excitation, its accurate treatment is known to be challenging, and consequently, cyanine dyes have often been used as model systems to assess the quality of electronic structure methods for excited states.\textsuperscript{4-12}

Here, we employ quantum Monte Carlo (QMC) to revisit the vertical excitation energies of cyanine dyes of the simple form $C_nH_{2n}(NH_2)_2^+$ with $n$ an odd number ranging from 1 to 17, combining the use of sophisticated multideterminant wave functions with recent developments for their efficient optimization in variational Monte Carlo (VMC).\textsuperscript{13-16} In particular, we build on our successful treatment at chemical accuracy of the excitation energies and optimal excited-state structures of small, prototypical molecules,\textsuperscript{17-19} where the determinantal components of the multiple states are generated in an automatic and balanced manner with the configuration interaction using a perturbative selection made iteratively (CIPSI) approach.\textsuperscript{20} Studying the bright excitation of cyanine dyes enables us to demonstrate the accuracy of our protocol for the shorter chains, where high-level coupled cluster (CC) offers a good compromise in terms of accuracy versus computational cost. Importantly, it also establishes the applicability of QMC to larger sizes where the use of other high-level approaches is more challenging. Finally, we identify the key descriptors of orbital correlations for these systems and elucidate why earlier QMC studies with limited active space wave functions lacked the expected accuracy.\textsuperscript{4}

2. METHODS

We employ QMC wave functions of the so-called Jastrow-Slater form, namely

$$\Psi = \mathcal{J} \sum_{i=1}^{N_{\text{eff}}} c_i D_i$$

(1)

where $\mathcal{J}$ is the Jastrow correlation factor, and $D_i$ are determinants of single-particle orbitals. The Jastrow factor explicitly depends on the interparticle coordinates and includes here electron–electron and electron–nucleus correlation terms.\textsuperscript{21}

Received: November 17, 2021
Published: January 26, 2022
To generate the determinantal components for the two states, we employ the CIPSI approach which, starting from a given reference space, builds expansions by iteratively selecting determinants based on their second-order perturbation (PT2) energy contribution obtained via the Epstein-Nesbet partitioning of the Hamiltonian \[ \delta E^{(2)}_\alpha = \frac{\langle \alpha | \hat{H} (|\Psi^{\text{CIPSI}}\rangle \rangle^2}{\langle \Psi^{\text{CIPSI}} | \hat{H} (|\Psi^{\text{CIPSI}}\rangle \rangle - \langle \alpha | \hat{H}(\alpha) \rangle } \] (2)

where $|\Psi^{\text{CIPSI}}\rangle$ is the current CIPSI wave function for the state under consideration, and $|\alpha\rangle$ denotes a determinant outside the current CI space. Since the ground and excited states of the cyanines have different symmetries, a state-specific approach can be used to perform the selection for the two states separately, using different orbitals.

We are here interested in computing excitation energies and, therefore, wish to achieve a balanced CIPSI description of the states of interest, which leads to converged excitation energies in QMC already for relatively small expansions. A measure of the quality of a given CIPSI wave function is its PT2 energy contribution, which represents an approximate estimate of the error of the expansion with respect to the full CI (FCI) limit.

Therefore, we compute the excitation energies using expansions for the two (or more generally multiple) states with matched PT2 energy and, therefore, ensure comparable quality. We refer the reader to ref 19 on how to impose the matched PT2 energy and, therefore, ensure comparable expansions for the two (or more generally multiple) states with different orbitals.

Alternatively, one can match the CI variance of the relevant states, which is defined as the variance of the FCI Hamiltonian:

\[ \sigma^2_C(|\Psi^{\text{CIPSI}}\rangle) = \sum_{i \in \text{FCI}} \langle \Psi^{\text{CIPSI}} | \hat{T} (|\Psi^{\text{CIPSI}}\rangle \rangle (\langle \Psi^{\text{CIPSI}} | \hat{T} (|\Psi^{\text{CIPSI}}\rangle \rangle) - \langle \Psi^{\text{CIPSI}} | \hat{T} (|\Psi^{\text{CIPSI}}\rangle \rangle)^2 = \sum_{\alpha} (\langle \Psi^{\text{CIPSI}} | \hat{T}(\alpha) \rangle)^2 \] (3)

As the CIPSI wave function approaches the FCI limit, the CI variance goes to zero. For various small molecules, we have found that matching the PT2 energy contributions leads to expansions with also very similar variances. In general, this is not always the case, and one of the two criteria might be more suitable than the other for the computation of the CI excitation energies of a particular system.

While we discuss in detail below the impact of this choice on the QMC excitation energies, we stress already here that the convergence of the QMC results is established not based on their agreement with available reference data but in an “internally consistent” manner based on the similarity of the VM and DMC excitation energies and their convergence with respect to the number of determinants.

Finally, as an alternative to the CIPSI expansions, we test complete active space (CAS) expansions for the determinantal components of our QMC wave functions. We start from separate CASSCF calculations for the two states and consider minimal active spaces by correlating the $\pi$ electrons in the $\pi$ orbitals constructed from the $2p_z$ orbitals. For the smaller cyanines with up to 7 heavy atoms, CN3–CN7 (we label a cyanine as CNm with $m$ the total number of C and N atoms), we also explore the use of a larger active space with molecular orbitals constructed from the $2p_z$ and $3p_z$ atomic orbitals. Finally, in some cases, we also test the performance of a simple one-configuration ansatz, namely, the Hartree–Fock (HF) and HOMO–LUMO (HL) configurations for the ground and excited states, respectively.

### 3. COMPUTATIONAL DETAILS

Unless otherwise specified, we employ scalar-relativistic energy-consistent HF pseudopotentials and the correlation-consistent Gaussian basis sets specifically constructed for these pseudopotentials. For most of the calculations, we use a double-$\zeta$ basis set minimally augmented with $s$ and $p$ diffuse functions on the heavy atoms and denoted here as maug-cc-pVDZ. Convergence tests are performed with the fully augmented aug-cc-pVTZ basis set. The exponents of the diffuse functions are taken from the corresponding all-electron Dunning’s correlation-consistent basis sets.

The HF and CASSCF computations are performed with the program GAMESS(US) When using the CASSCF wave functions in QMC, we truncate the CAS expansion for CN11 and CN13, using a threshold on the CSF coefficients so that the configurations make up respectively about 0.9985 and 0.9765 of the weight of the total wave functions of the two states. The CIPSI expansions are generated with Quantum Package and constructed to be eigenstates of $\hat{S}_z^2$. We perform the selection for the two states separately, starting from CASSCF orbitals obtained with the larger active spaces for the cyanine molecules up to CN7 and the minimal CAS from CN9 to CN15. We use the HF orbitals for CN17 and CN19. As shown in Figure S1 and Table S5 for CN3 and Figure S6 for CN15, the use of different orbitals to generate the CIPSI expansions has no appreciable impact on the CI or QMC excitation energies.

The QMC calculations are carried out with the CHAMP code. The determinantal part of our QMC wave functions is expressed in terms of spin-adapted configuration state functions (CSF) to reduce the number of parameters during the VMC optimization. In the wave function optimization, we sample a guiding wave function that differs from the current wave function close to the nodes to guarantee finite variances of the estimators of the gradients with respect to the wave function parameters. All wave function parameters (Jastrow, CI, and orbital coefficients) are optimized in state-specific energy minimization following the stochastic reconfiguration scheme. In the DMC calculations, we treat the pseudopotentials beyond the locality approximation using the T-move algorithm and employ an imaginary time-step of 0.05 au which we have already tested for one of the cyanine chains and shown to yield excitation energies converged to better than 0.01 eV.

We compute all energies on the ground-state geometries of CN3–CN11 determined with all-electron PBE0/cc-pVQZ in ref 8 and obtain the geometries for CN13 to CN19 at the same level of theory with the Gaussian 09 program. We employ the programs CFour v2.135 and Molcas 36 for the approximate coupled cluster singles and doubles (CC2) and singles, doubles, and triples models (CC3) and the CASPT2 calculations, respectively, using the all-electron aug-cc-pVDZ basis set and the frozen-core approximation, unless otherwise specified.

### 4. RESULTS

We compute the lowest $\pi \rightarrow \pi^*$ vertical excitation energy of cyanine dyes of the form $C_nH_{2n}(NH_2)^\pi$ with $n$ ranging from 1 up to 17. The structures of the CN3 and CN9 molecules are
shown in Figure 1. In all cases, the point group of the molecule is \( C_{2v} \), with the ground (GS) and excited (ES) states having \( A_{1} \) and \( B_{1} \) symmetry, respectively.

For CN3 up to CN15, we compare the QMC excitation energies with the all-electron CC3/aug-cc-pVDZ results. The use of the CC3 method as reference for the bright excitation of these systems is supported by the agreement of the CC3 excitation energies with the corresponding extrapolated FCI (exFCI) estimates in a small basis of the smaller CN3 and CN5 to better than 0.05 eV. Employing the aug-cc-pVDZ basis set is sufficient given the agreement with the corresponding aug-cc-pVTZ values (see Table S1). Importantly, the all-electron CC3/aug-cc-pVDZ excitation energies are very close to the BFD CC3/aug-cc-pVDZ values, confirming that the use of pseudopotentials does not introduce appreciable errors. The reference CC3/aug-cc-pVDZ values also agree with the corresponding CC3 excitation energies computed with the BFD maug-cc-pVDZ basis set for all cyanines except the smallest CN3 (see Table S1), where a fully augmented double-\( \zeta \) basis is needed also in the BFD calculations.

For dyes larger than CN15, we are however not able to run the CC3 calculations due to memory requirements, and the DMC excitation energy with our best CIPSI wave function becomes then the reference for other calculations.

4.1. Building the Expansions. To compute accurate QMC excitation energies for the cyanine dyes, one needs balanced Jastrow-Slater wave functions to describe the ground- and excited-state CIPSI expansions having either matched PT2 energy contributions or CI variances. The VMC and DMC excitation energies obtained using the iso-variance expansions are also shown (the statistical error is smaller than the symbol size). The BFD pseudopotentials and the maug-cc-pVDZ basis are used here also for the CC3 calculations.

Figure 1. CI vertical excitation energies of CN3 (top) and CN9 (bottom) versus the total number of determinants, computed for ground- and excited-state CIPSI expansions having either matched PT2 energy contributions or CI variances. The VMC and DMC excitation energies obtained using the iso-variance expansions are also shown (the statistical error is smaller than the symbol size). The BFD pseudopotentials and the maug-cc-pVDZ basis are used here also for the CC3 calculations.

![Figure 1](image)

Table 1. Vertical Excitation Energies (eV) for the Cyanine Dyes Computed with QMC and Other Highly-Correlated Methods

| method                  | CN3 | CN5 | CN7 | CN9 | CN11 | CN13 | CN15 | CN17 | CN19 |
|-------------------------|-----|-----|-----|-----|------|------|------|------|------|
| VMC-CAS                 | 7.67(1) | 5.13(1) | 3.97(1) | 3.11(1) | 2.58(1) | 2.13(1) |
| DMC-CAS                 | 7.49(1) | 5.04(1) | 3.83(1) | 3.04(1) | 2.55(1) | 2.15(1) |
| VMC-CIPSI               | 7.23(1) | 4.83(1) | 3.65(1) | 3.03(1) | 2.55(1) | 2.18(1) | 1.85(1) | 1.66(1) | 1.59(2) |
| DMC-CIPSI               | 7.23(1) | 4.86(1) | 3.66(1) | 2.98(1) | 2.54(1) | 2.15(1) | 1.90(1) | 1.65(1) | 1.57(1) |
| CASPT2/aug-cc-pVDZ      | 6.94 | 4.64 | 3.56 | 2.91 | 2.45 | 2.11 | 1.85 | 1.65 |
| CC2/aug-cc-oVDZ         | 7.29 | 4.97 | 3.80 | 3.10 | 2.64 | 2.30 | 2.04 | 1.84 |
| CC3/aug-cc-pVDZ         | 7.20 | 4.85 | 3.67 | 2.97 | 2.50 | 2.16 | 1.91 |
| exFCI/aug-cc-pVDZ\(^b\) | 7.17(2) | 4.89(2) |

\(^a\)The QMC-CIPSI calculations for CN3 are performed with the aug-cc-pVTZ basis. \(^b\)The BFD pseudopotentials are used in QMC, while all other calculations are all-electron. All energies are computed on PBE0/cc-pVQZ geometries.
the CI treatment of the smallest cyanine, CN3, appears to be the most difficult as further elaborated in section 4.3.

Importantly, in Figure 1, we also show that QMC largely corrects for possible shortcomings of the starting CIPSI expansions, yielding excitation energies which display a rather small dependence on the number of determinants, especially at the DMC level. For CN3 and small expansions, where the iso-variance criterion significantly overestimates the CI excitation energy, VMC and DMC reduce the error at the CI level by about 0.2 and 0.3 eV, respectively. As the expansions become larger, the difference between the VMC and the DMC values diminishes, falling well below chemical accuracy (about 0.05 eV) for both CN3 and CN9. The robustness of the QMC results is further corroborated for CN7 in Table S6, where we show that, for a comparable number of determinants, the use of PT2- and variance-matched wave functions yields excitation energies which differ by about 0.2 eV at the CI level but are very close in VMC and completely equivalent in DMC. 4.2. Best QMC Vertical Excitations. In Table 1, we summarize the VMC and DMC excitation energies of all cyanine dyes obtained with the largest CIPSI expansions of Table S6 and the iso-variance selection criterion. We also list the QMC and CASPT2 excitation energies computed with minimal CAS expansions, together with our CC2 and CC3 results and the exFCI estimates from the literature.12 We refer the reader to Table S6 for additional QMC calculations with different numbers of determinants in the Jastrow-CIPSI wave functions.

For the reported CIPSI expansions, the VMC and DMC excitation energies are very close and also agree within chemical accuracy with the CC3 and exFCI values in all cases where these methods are applicable. This is in line with our previous findings that the agreement between VMC and DMC excitation energies is a strong indication of the balanced quality of the corresponding wave functions.19 Furthermore, we find that the QMC values for the larger dyes are in very good agreement with the estimates given by the extrapolation of the CC3 results as a function of the number of electrons (see the SI). Since DMC can be employed in all cases, we plot all excitation energies in Figure 2 in terms of their distance to the DMC-CIPSI results, which we use as reference values.

The CASPT2 and QMC-CAS energies computed with the minimal active spaces are instead very different from the DMC-CIPSI results: CASPT2 always underestimates the excitation energies, whereas QMC-CAS tends to overestimate them, similarly to what was reported for CAS wave functions in ref 4. For CN3—CN7, we test the effect of including more π orbitals in the active space, which somewhat ameliorates the VMC excitation energies but does not sufficiently affect the DMC values, which remain far from the DMC-CIPSI reference (see Table S7). Interestingly, we note that both the CASPT2 and QMC-CAS methods approach the best DMC results as the size of the molecule increases, suggesting an easier treatment of the longer chains as already found at the CI level and as further discussed below.

4.3. Capturing Orbital Correlation. To understand the different performance of CAS and CIPSI expansions when used in QMC wave functions, we focus here on CN3 and analyze in Figure 3 the VMC and DMC vertical excitation energies calculated using different determinantal components in the trial wave functions. As already mentioned, despite being the smallest cyanine dye, CN3 appears to be the most challenging one: the use of Jastrow-CAS wave functions leads to quite big errors, and the number of CIPSI determinants needed to converge the excitation energy is larger than for the longer dyes.

The simplest QMC calculations are performed with a one-configuration (HF/HL) wave function and the maug-cc-pVDZ basis set. We then proceed to CAS determinantal components and CIPSI expansions also employing the aug-cc-pVTZ basis set. The VMC excitation energy computed with the minimal CAS wave functions is worse than the HF/HL value since the active space comprises more determinants for the ground state but only the HL configuration for the excited state. DMC ameliorates the result, but using a larger CAS space on the π orbitals only marginally helps (see Table S7). On the other hand, with the CIPSI selected determinants, we have a considerable improvement on the excitation energy, and with the use of just a few hundred determinants, the DMC error reduces to less than 0.1 eV. Employing larger expansions with the maug-cc-pVDZ basis set, we finally converge to VMC values which are consistent with the DMC ones and approximately 0.04 eV higher than the reference. The use of the aug-cc-pVTZ basis set further reduces the excitation energy by about 0.02 eV. We note that, for the longer cyanine chains, the smaller maug-cc-pVDZ basis set is found to be sufficient for the computation of this excitation energy.18

The superior performance of the use of a CIPSI with respect to the CAS expansions in QMC indicates that some key
We correlate each state, we compute the energy with only one CSF, and on only the valence electrons and use state-average natural orbitals. The reference FCI excitation energy in this basis as in ref 4, the active space is chosen to correlate the π electrons in the π orbitals. From the QMC-CIPSI results, we can therefore infer that, while the excitation of interest is predominantly of π → π* character, other orbital correlations are important and cannot be omitted in the QMC wave function of the shorter cyanines.

To better understand this, we present a CI study for CN3–CN7 with the small 6-31G basis set in Table 2. We correlate only the valence electrons and use state-average natural orbitals obtained with a preliminary calculation at the CIPSI level. For each state, we compute the energy with only one CSF, and on top of this configuration, we perform a CAS-CI calculation restricted to the σ and the π orbitals in a CAS-σ and CAS-π, respectively. The reference FCI excitation energy in this basis and the associated confidence interval are computed following the scheme presented in ref 38 rather than extrapolating the variational energies of the individual states in the limit of the PT2 energy correction going to zero. Indeed, the uncertainties of the extrapolated FCI energies of both states are larger than the uncertainty on the estimated excitation energy computed with this scheme. Since the CC3 estimate for CN3 and CN5 is in excellent agreement with the FCI value, we use the CC3 excitation energy as reference for CN7. We note that, because of the use of the simple 6-31G basis set, the FCI and CC3 excitation energies are much higher than the more accurate results presented above, but this is not relevant for the present discussion.

For the CN3 molecule, the excitation energy obtained with a single CSF for each state is 8.91 eV, namely, higher by 1.4 eV than the FCI result. The CAS-π calculation corrects only 60% of the error, indicating that the σ orbitals also play an important role in the stabilization of the excited state. Similarly, the excitation energy obtained with the CAS-σ improves the excitation energy with respect to the single CSF by recovering about 31% of the error. These results indicate the importance of both σ and π orbitals in the calculation of the excitation energy of CN3.

Therefore, to partially account for both π and σ correlations, we perform a multireference CI calculation, applying all possible single and double excitations to the CAS-π determinants. Such a CAS-π+SD-σ calculation also enables the relaxation of the CAS-π CI coefficients in the presence of most of the σ correlation. The resulting excitation energy is now significantly improved but still 0.1 eV higher than the FCI reference, confirming that a similar computational effort needs to be made for the π and σ orbitals. This justifies the use of CIPSI where the most important Slater determinants will be chosen to describe σ, π, and σ→π correlation in a “democratic” way based on their contribution to the second-order perturbation energy.

For CN5, the situation is somewhat different. While the single CSF still overestimates the excitation energy by 1.47 eV, the CAS-π wave function behaves better than for CN3, recovering 80% of the error. Consequently, omitting the σ orbitals in the active space results in an excitation energy closer to the reference than in the CN3 case. Once the σ orbitals are introduced, as for CN3, we improve the excitation energy but still observe an overestimation of the CAS-π+SD-σ result by almost 0.2 eV, pointing to the importance of describing the σ as well as the π correlation. The situation for CN7 is similar as for CN5, suggesting that the whole series behaves like CN5 and that CN3 is an exception because of the particularly small length of the chain.

### 5. CONCLUSION

We have presented a QMC benchmark study of the lowest vertical excitation energies of cyanine chains. We constructed the determinantal components of the Jastrow-Slater wave functions through an automatic selected-CI procedure and obtained a balanced description of the relevant states by ensuring similar quality of the corresponding expansions, for instance by matching their CI variances. With compact descriptor of correlation is missing from the active space and is not recovered through the addition of the Jastrow factor and the subsequent full optimization in VMC nor through a DMC calculation with the optimal Jastrow-Slater wave function. In this work as in ref 4, the active space is chosen to correlate the π electrons in the π orbitals. From the QMC-CIPSI results, we can therefore infer that, while the excitation of interest is predominantly of π → π* character, other orbital correlations are important and cannot be omitted in the QMC wave function of the shorter cyanines.

To better understand this, we present a CI study for CN3–CN7 with the small 6-31G basis set in Table 2. We correlate only the valence electrons and use state-average natural orbitals obtained with a preliminary calculation at the CIPSI level. For each state, we compute the energy with only one CSF, and on top of this configuration, we perform a CAS-CI calculation restricted to the σ and the π orbitals in a CAS-σ and CAS-π, respectively. The reference FCI excitation energy in this basis and the associated confidence interval are computed following the scheme presented in ref 38 rather than extrapolating the variational energies of the individual states in the limit of the PT2 energy correction going to zero. Indeed, the uncertainties of the extrapolated FCI energies of both states are larger than the uncertainty on the estimated excitation energy computed with this scheme. Since the CC3 estimate for CN3 and CN5 is in excellent agreement with the FCI value, we use the CC3 excitation energy as reference for CN7. We note that, because of the use of the simple 6-31G basis set, the FCI and CC3 excitation energies are much higher than the more accurate results presented above, but this is not relevant for the present discussion.

For the CN3 molecule, the excitation energy obtained with a single CSF for each state is 8.91 eV, namely, higher by 1.4 eV than the FCI result. The CAS-π calculation corrects only 60% of the error, indicating that the σ orbitals also play an important role in the stabilization of the excited state. Similarly, the excitation energy obtained with the CAS-σ improves the excitation energy with respect to the single CSF by recovering about 31% of the error. These results indicate the importance of both σ and π orbitals in the calculation of the excitation energy of CN3.

Therefore, to partially account for both π and σ correlations, we perform a multireference CI calculation, applying all possible single and double excitations to the CAS-π determinants. Such a CAS-π+SD-σ calculation also enables the relaxation of the CAS-π CI coefficients in the presence of most of the σ correlation. The resulting excitation energy is now significantly improved but still 0.1 eV higher than the FCI reference, confirming that a similar computational effort needs to be made for the π and σ orbitals. This justifies the use of CIPSI where the most important Slater determinants will be chosen to describe σ, π, and σ→π correlation in a “democratic” way based on their contribution to the second-order perturbation energy.

For CN5, the situation is somewhat different. While the single CSF still overestimates the excitation energy by 1.47 eV, the CAS-π wave function behaves better than for CN3, recovering 80% of the error. Consequently, omitting the σ orbitals in the active space results in an excitation energy closer to the reference than in the CN3 case. Once the σ orbitals are introduced, as for CN3, we improve the excitation energy but still observe an overestimation of the CAS-π+SD-σ result by almost 0.2 eV, pointing to the importance of describing the σ as well as the π correlation. The situation for CN7 is similar as for CN5, suggesting that the whole series behaves like CN5 and that CN3 is an exception because of the particularly small length of the chain.
expansions of only a few thousand determinants, upon
optimization of all parameters in our wave functions, we
obtained QMC excitation energies which improve on the
starting CI values and, for the shorter chain lengths where CC3
calculations are feasible, agree with the CC3 results to chemical
accuracy. We also applied our protocol to longer cyanines and
validated the accuracy of our estimates via the consistent
closeness of the determined VMC and DMC excitation
energies. Finally, we showed that key to a successful
description of this excitation over all chain lengths is to
account for π, σ, and σ−π correlations, therefore going beyond a
CAS treatment based on σ-orbitals only. In conclusion, we
believe that the present study further establishes QMC
methods as accurate and robust tools for the treatment of
excited states of relatively large systems and parameter spaces.

**ASSOCIATED CONTENT**

*Supporting Information*

The Supporting Information is available free of charge at
https://pubs.acs.org/10.1021/acs.jctc.1c01162.

Dependence of excitation energies on basis set and
pseudopotential; detailed information on all CIPSI and
QMC calculations for CN3; CIPSI energies and excitation
energies obtained by matching either PT2
energy or CI variance for all dyes; impact of using HF
orbitals in generating CIPSI expansions; QMC-CIPSI and
QMC-CAS energies and excitation energies; and
PBE0/cc-pVQZ geometries of CN13-CN19 (PDF)

**AUTHOR INFORMATION**

**Corresponding Authors**

Saverio Moroni — CNR-IOM DEMOCRITOS, Istituto
Officina dei Materiali, and SISSA Scuola Internazionale
Superiore di Studi Avanzati, I-34136 Trieste, Italy;
Email: moroni@democritos.it

Anthony Scemama — Laboratoire de Chimie et Physique
Quantiques, Université de Toulouse, CNRS, UPS, 31062
Toulouse, France; orcid.org/0000-0003-4955-7136;
Email: scemama@irsamc.ups-tlse.fr

Claudia Filippi — MESA+ Institute for Nanotechnology,
University of Twente, 7500 AE Enschede, The Netherlands;
orcid.org/0000-0002-2425-6735; Email: c.filippi@utwente.nl

**Author**

Alice Cuzzocrea — MESA+ Institute for Nanotechnology,
University of Twente, 7500 AE Enschede, The Netherlands;
orcid.org/0000-0001-7446-9643

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c01162

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

A.C. is supported by the “Computational Science for Energy
Research and Netherlands eScience Center joint program”
(project CSER,JCEL.022) of The Netherlands Organisation for
Scientific Research (NWO). The calculations were carried
out on the Dutch national supercomputer Cartesius with the
support of SURF Cooperative. The work is also partially
supported by the European Center of Excellence in Exascale
Computing TREX - Targeting Real Chemical Accuracy at the
Exascale, funded from the European Union’s Horizon 2020
Research and Innovation program (grant no. 952165).

**REFERENCES**

(1) Levitus, M.; Ranjit, S. Cyanine dyes in biophysical research: the
photophysics of polymethylene fluorescent dyes in biomolecular
environments. *Q. Rev. Biophys.* 2011, **44**, 123–151.

(2) Shindy, H. Fundamentals in the chemistry of cyanine dyes: A
review. *Dyes Pigm.* 2017, **145**, S05–S13.

(3) Oprea, C. I.; Panait, P.; Essam, Z. M.; Abd El-Aal, R. M.; Girțu,
M. A. Photoexcitation Processes in Oligomethylene Cyanine Dyes for
Dye-Sensitized Solar Cells—Synthesis and Computational Study.
*Nanomaterials* 2020, **10**, 662.

(4) Send, R.; Valsson, O.; Filippi, C. Electronic Excitations of Simple
Cyanine Dyes: Reconciling Density Functional and Wave Function
Methods. *J. Chem. Theory Comput.* 2011, **7**, 444–455.

(5) Jacquemin, D.; Zhao, Z.; Valero, R.; Adamo, C.; Ciocini, I.;
Truhlar, D. Verdict: Time-Dependent Density Functional Theory
“Not Guilty” of Large Errors for Cyanines. *J. Chem. Theory Comput.*
2012, **8**, 1255–1259.

(6) Moore, B.; Autschbach, J. Longest-Wavelength Electronic
Excitations of Linear Cyanines: The Role of Electron Delocalization
and of Approximations in Time-Dependent Density Functional Theory.
*J. Chem. Theory Comput.* 2013, **9**, 4991–5003.

(7) Zheko, H.; Kryukov, M.; Autschbach, J.; Ziegler, T. Applications of
Time Dependent and Time Independent Density Functional Theory to the
First π to π* Transition in Cyanine Dyes. *J. Chem. Theory Comput.*
2014, **10**, 3299–3307.

(8) Boulanger, P.; Jacquemin, D.; Duchemin, I.; Blase, X. Fast and
Accurate Electronic Excitations in Cyanines with the Many-Body
Bethe-Slater Approach. *J. Chem. Theory Comput.* 2014, **10**, 1212–
1218.

(9) Filatov, M.; Huix-Rotllant, M. Assessment of density functional
too
theory based ΔSCF (self-consistent field) and linear response
methods for longest wavelength excited states of extended π-
conjugated molecular systems. *J. Chem. Phys.* 2014, **141**, 024112.

(10) Le Guennic, B.; Jacquemin, D. Taking Up the Cyanine
Challenge with Quantum Tools. *Acc. Chem. Res.* 2015, **48**, 530–537.

(11) Minezawa, N. Vertical excitation energies of linear cyanine dyes
by spin-flip time-dependent density functional theory. *Chem. Phys.
Lett.* 2015, **622**, 115–119.

(12) Garniron, Y.; Anthony, S.; Emmanuel, G.; Michel, C.; Loos, P.-F.
Selected configuration interaction interaction dressed by perturbation.
*J. Chem. Phys.* 2018, **149**, 064103.

(13) Sorrella, S.; Capriotti, L. Algorithmic differentiation and the
calculation of forces by quantum Monte Carlo. *J. Chem. Phys.*
2010, **133**, 234111.

(14) Neuscamman, E.; Umrigar, C. J.; Chan, G. K.-L. Optimizing
large parameter sets in variational quantum Monte Carlo. *Phys. Rev. B*
2012, **85**, 045103.

(15) Filippi, C.; Assaraf, R.; Moroni, S. Simple formalism for efficient
derivatives and multi-determinant expansions in quantum Monte
Carlo. *J. Chem. Phys.* 2016, **144**, 194105.

(16) Assaraf, R.; Moroni, S.; Filippi, C. Optimizing the Energy with
Quantum Monte Carlo: A Lower Numerical Scaling for Jastrow−
 Slater Expansions. *J. Chem. Theory Comput.* 2017, **13**, 5273–5281.

(17) Dash, M.; Feldt, J.; Moroni, S.; Scemama, A.; Filippi, C. Excited
States with Selected Configuration Interaction-Quantum Monte
Carlo: Chemically Accurate Excitation Energies and Geometries.
*J. Chem. Theory Comput.* 2019, **15**, 4896–4906.

(18) Cuzzocrea, A.; Scemama, A.; Briels, W. J.; Moroni, S.; Filippi,
C. Variational Principles in Quantum Monte Carlo: The Troubled
Story of Variance Minimization. *J. Chem. Theory Comput.* 2020, **16**, 4203–4212.

(19) Dash, M.; Moroni, S.; Filippi, C.; Scemama, A. Tailoring CIPSI
Expansions for QMC Calculations of Electronic Excitations: The Case
Study of Thiophene. *J. Chem. Theory Comput.* 2021, **17**, 3426–3434.
(20) Huron, B.; Malrieu, J.; Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. J. Chem. Phys. 1973, 58, 5745–5759.

(21) As the Jastrow factor, we use the exponential of the sum of two fifth-order polynomials of the electron–nuclear and the electron–electron distances, respectively, and rescale the interparticle distances as $R = (1 - \exp(-\kappa r))/x$ with $x$ set to 0.6 au. We employ different electron-nucleus Jastrow factors to describe the correlation of an electron with C, N, and H. The total number of free parameters to be optimized in the Jastrow factor is 17 for the systems considered here.

(22) Epstein, P. S. The stark effect from the point of view of Schrödinger’s quantum theory. Phys. Rev. 1926, 28, 695.

(23) Nesbet, R. K. Configuration interaction in orbital theories. Proc. R. Soc. London A 1955, 230, 312–321.

(24) Burkatzi, M.; Filippi, C.; Dolg, M. Energy-consistent pseudopotentials for quantum Monte Carlo calculations. J. Chem. Phys. 2007, 126, 234105.

(25) For the hydrogen atom, we use a more accurate BFD pseudopotential and basis set. Dolg, M.; Filippi, C. private communication.

(26) Kendall, R. A.; Dunning, T. H., Jr; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. J. Chem. Phys. 1992, 96, 6796–6806.

(27) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; others. General atomic and molecular electronic structure system. J. Comput. Chem. 1993, 14, 1347–1363.

(28) Gordon, M. S.; Schmidt, M. W. Theory and applications of computational chemistry; Elsvier: 2005; pp 1167–1189.

(29) Garniron, Y.; Applencourt, C.; Gasperich, B.; Benali, A.; Ferté, A.; Paquier, J.; Pradines, B.; Renault, P.; Toulouse, J.; Barbareoso, P.; Renon, N.; David, G.; Malrieu, J.-P.; Veril, M.; Caffarel, M.; Loos, P.-F.; Giner, E.; Scemama, A. Quantum Package 2.0: An Open-Source Determinant-Driven Suite of Programs. J. Chem. Theory Comput. 2019, 15, 3591–3609.

(30) CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, S. Moroni, and collaborators.

(31) Attaccalite, C.; Sorella, S. Stable liquid hydrogen at high pressure by a novel ab initio molecular-dynamics calculation. Phys. Rev. Lett. 2008, 100, 114501.

(32) Sorella, S.; Casula, M.; Rocca, D. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. J. Chem. Phys. 2007, 127, 014105.

(33) Casula, M. Beyond the locality approximation in the standard diffusion Monte Carlo method. Phys. Rev. B 2006, 74, 161102.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Stratmann, R. E.; Farkas, O.; Tomasi, J.; Ochterski, J. W.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Millam, J. M.; Klene, M.; Adams, M.; Zakrzewski, V. G.; Li, X.; Piskorz, P.; Komaromi, I.; Nanninga, J.; Gonzalez, C.; Head-Gordon, M.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. J.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 09, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2016.

(35) Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G. CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package. With contributions from Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Blaschke, S.; Bomble, Y. J.; Burger, S.; Christiansen, O.; Datta, D.; Engel, F.; Faber, R.; Greiner, J.; Heckert, M.; Heun, O.; Hilgenberg, M.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Jesulius, J.; Kirsch, T.; Klein, K.; Kopper, G. M.; Launderdale, W. J.; Lipparrini, F.; Metzroth, T.; Müller, L. A.; O’Neill, D. P.; Nottoli, T.; Price, D. R.; Prochnow, E.; Puzzarini, C.; Ruud, K.; Schiﬀmann, F.; Schwabach, W.; Simmons, C.; Stopkowicz, S.; Tajti, A.; Vázquez, J.; Wang, F.; Watts, J. D. and the integral packages MOLECULE (Amlój, J.; Taylor, P. R.), PROPS (Taylor, P. R.), ABACUS (Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J.), and ECP routines by Mitin, A. V.; van Wüllen, C. For the current version, see: http://www.cfour.de (accessed 2022-01-21).

(36) Aquilante, F.; Autschbach, J.; Carlsson, R. K.; Chibotaru, L. F.; Delcey, M. I. G.; De Vico, L.; Fdez; Galván, I.; Ferré, N.; Frutos, L. M.; Gagliardi, L.; Garavelli, M.; Giussani, A.; Hoyer, C. E.; Li Manni, G.; Lischka, H.; Ma, D.; Malmqvist, P. r.; Müller, T.; Nenov, A.; Olivucci, M.; Pedersen, T. B.; Peng, D.; Plasser, F.; Pritchard, B.; Reiher, M.; Rivalta, I.; Schapiro, J.; Segarra-Martí, J.; Sterup, M.; Truhlar, D. G.; Ungur, L.; Valentin, A.; Vanzo, S.; Veryazov, V.; Vysotskiy, V. P.; Weingart, O.; Zapata, F.; Lindh, R. Molcas 8: New capabilities for multiconfigurational quantum chemical calculations across the periodic table. J. Comput. Chem. 2016, 37, 506–541.

(37) The CC calculations performed with Cfour code are run on a dual socket AMD Epyc 7402 with 256GB of memory.

(38) Véril, M.; Scemama, A.; Caffarel, M.; Lipparrini, F.; Boggio-Pasqua, M.; Jacquemin, D.; Loos, P.-F. QUESTDB: A database of highly accurate excitation energies for the electronic structure community. WIREs Comput. Mol. Sci. 2021, 11, e1517.