The Journal of Chemical Physics

The performance of CIPSI on the ground state electronic energy of benzene

Pierre-François Loos [iD], **Yann Damour, and Anthony Scemama** [iD]

The performance of CIPSI on the ground state electronic energy of benzene

View Online   Export Citation   CrossMark

Pierre-François Loos,[a] [iD] Yann Damour, and Anthony Scemama[b] [iD]

AFFILIATIONS

Laboratoire de Chimie et Physique Quantiques (UMR 5626), Université de Toulouse, CNRS, UPS, Toulouse, France

[a] Author to whom correspondence should be addressed: loos@irsamc.ups-tlse.fr
[b] scemama@irsamc.ups-tlse.fr

ABSTRACT

Following the recent work of Eriksen *et al.* [J. Phys. Chem. Lett. **11**, 8922 (2020)], we report the performance of the *configuration interaction using a perturbative selection made iteratively* method on the non-relativistic frozen-core correlation energy of the benzene molecule in the cc-pVDZ basis. Following our usual protocol, we obtain a correlation energy of -863.4 mE_h, which agrees with the theoretical estimate of -863 mE_h proposed by Eriksen *et al.* [J. Phys. Chem. Lett. **11**, 8922 (2020)] using an extensive array of highly accurate new electronic structure methods.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0027617

Although sometimes decried, one cannot deny the usefulness of benchmark sets and their corresponding reference data for the electronic structure community. These are, indeed, essential for the validation of existing theoretical models and to bring to light and subsequently understand their strengths and, more importantly, their weaknesses. In that regard, the previous benchmark datasets provided by the *Simons Collaboration on the Many-Electron Problem* have been extremely valuable.[1–3] The same comment applies to the excited-state benchmark set of Thiel and co-workers.[4–8] Following a similar goal, we have recently proposed a large set of highly accurate vertical transition energies for various types of excited states, thanks to the renaissance of selected configuration interaction (SCI) methods,[9–11] which can now routinely produce near full configuration interaction (FCI) quality excitation energies for small- and medium-sized organic molecules.[12–16]

In a recent article,[17] Eriksen *et al.* have proposed a blind test for a particular electronic structure problem inviting several groups around the world to contribute to this endeavor. In addition to coupled cluster theory with singles, doubles, triples, and quadruples (CCSDTQ),[18,19] a large panel of highly accurate, emerging electronic structure methods were considered: (i) the many-body expansion FCI (MBE-FCI),[20–23] (ii) three SCI methods including a second-order perturbative (PT2) correction [adaptive sampling CI (ASCI),[24–26] iterative CI (iCI),[27–30] and semistochastic heat-bath CI (SHCI)[31–33]], (iii) the full coupled-cluster reduction (FCCR)[34,35]

that also includes a second-order perturbative correction, (iv) the density-matrix renormalization group (DMRG) approach,[36–38] and (v) two flavors of FCI quantum Monte Carlo (FCIQMC),[39,40] namely, AS-FCIQMC[41] and CAD-FCIQMC.[42,43] We refer the interested reader to Ref. 17 and its supplementary material for additional details on each method and the complete list of references. Soon after, Lee *et al.* reported phaseless auxiliary-field quantum Monte Carlo[42,44] (ph-AFQMC) correlation energies for the very same problem.[45]

The target application is the non-relativistic frozen-core correlation energy of the ground state of the benzene molecule in the cc-pVDZ basis. The geometry of benzene has been optimized at the MP2/6-31G* level,[5] and its coordinates can be found in the supplementary material of Ref. 17 alongside its nuclear repulsion and Hartree–Fock energies. This corresponds to an active space of 30 electrons and 108 orbitals, i.e., the Hilbert space is of the order of 10^{35} Slater determinants. Needless to say, this size of Hilbert space cannot be tackled by exact diagonalization with the current architectures. The correlation energies reported in Ref. 17 are gathered in Table I alongside the best ph-AFQMC estimate from Ref. 45 based on a CAS(6,6) trial wave function. The outcome of this work is nicely summarized in the abstract of Ref. 17: "*In our assessment, the evaluated high-level methods are all found to qualitatively agree on a final correlation energy, with most methods yielding an estimate of the FCI value around -863 mE_h. However, we find the root-mean-square*

TABLE I. The frozen-core correlation energy ΔE (in mE_h) of benzene in the cc-pVDZ basis set using various methods.

Method	ΔE	References
ASCI	−860.0	17
iCI	−861.1	17
CCSDTQ	−862.4	17
DMRG	−862.8	17
FCCR	−863.0	17
MBE-FCI	−863.0	17
CAD-FCIQMC	−863.4	17
AS-FCIQMC	−863.7	17
SHCI	−864.2	17
ph-AFQMC	−864.3(4)	45
CIPSI	−863.4	This work

deviation of the energies from the studied methods to be considerable (1.3 mE_h), which in light of the acclaimed performance of each of the methods for smaller molecular systems clearly displays the challenges faced in extending reliable, near-exact correlation methods to larger systems."

For the sake of completeness and our very own curiosity, we report in this Note the frozen-core correlation energy obtained with a fourth flavor of SCI known as *Configuration Interaction using a Perturbative Selection made Iteratively* (CIPSI) method,[10] which also includes a PT2 correction. In short, the CIPSI algorithm belongs to the family of SCI+PT2 methods. The idea behind such methods is to slow down the exponential increase in the size of the CI expansion by retaining the most energetically relevant determinants only, thanks to the use of a second-order energetic criterion to select perturbatively determinants in the FCI space. However, performing SCI calculations rapidly becomes extremely tedious when one increases the system size as one hits the exponential wall inherently linked to these methods.

From a historical point of view, CIPSI is probably one of the oldest SCI algorithm. It was developed in 1973 by Huron, Rancurel, and Malrieu[10] (see also Ref. [46]). Recently, the determinant-driven CIPSI algorithm has been efficiently implemented[47,48] in the open-source programming environment QUANTUM PACKAGE by our group, enabling to perform massively parallel computations.[49–51] In particular, we were able to compute highly accurate ground- and excited-state energies for small- and medium-sized molecules (including benzene).[12–16] CIPSI is also frequently used to provide an accurate trial wave function for QMC calculations.[47,48,52–61] The particularity of the current implementation is that the selection step and the PT2 correction are computed *simultaneously* via a hybrid semistochastic algorithm[49,51] (which explains the statistical error associated with the PT2 correction in the following). Moreover, a renormalized version of the PT2 correction (dubbed rPT2 below) has been recently implemented and tested for a more efficient extrapolation to the FCI limit, thanks to a partial resummation of the higher-order of perturbation.[51] We refer the interested reader to Ref. [51] where one can find all the details regarding the implementation of the rPT2 correction and the CIPSI algorithm.

Being late to the party, we obviously cannot report blindly our CIPSI results. However, following the philosophy of Eriksen *et al.*[17] and Lee *et al.*,[45] we will report our results with the most neutral tone, leaving the freedom to the reader to make up his/her mind. We then follow our usual "protocol"[12–16,57–59] by performing a preliminary SCI calculation using Hartree–Fock orbitals in order to generate a SCI wave function with at least 10^7 determinants. Natural orbitals are then computed based on this wave function, and a new SCI calculation is performed with this new natural set of orbitals. This has the advantage to produce a smoother and faster convergence of the SCI energy toward the FCI limit. The total SCI energy is defined as the sum of the variational energy $E_{var.}$ (computed via diagonalization of the CI matrix in the reference space) and a second-order perturbative correction $E_{(r)PT2}$, which takes into account the external determinants, i.e., the determinants which do not belong to the variational space but are linked to the reference space via a nonzero matrix element. The magnitude of $E_{(r)PT2}$ provides a qualitative idea of the "distance" to the FCI limit. As mentioned above, SCI+PT2 methods rely heavily on extrapolation, especially when one deals with medium-sized systems. We then linearly extrapolate the total SCI energy to $E_{(r)PT2} = 0$ (which effectively corresponds to the FCI limit). Note that, unlike excited-state calculations where it is important to enforce that the wave functions are eigenfunctions of the $\hat{S}^2$ spin operator,[62] the present wave functions do not fulfill this property as we aim for the lowest possible energy of a singlet state. We have found that $\langle \hat{S}^2 \rangle$ is, nonetheless, very close to zero (~5×10^{-3} a.u.). The corresponding energies are reported in Table II as functions of the number of determinants in the variational space N_{det}.

A second run has been performed with localized orbitals. Starting from the same natural orbitals, a Boys–Foster localization procedure[63] was performed in several orbital windows: (i) core, (ii) valence σ, (iii) valence π, (iv) valence π^*, (v) valence σ^*, (vi) the higher-lying σ orbitals, and (vii) the higher-lying π orbitals.[64] Like Pipek–Mezey,[65] this choice of orbital windows allows us to preserve a strict σ–π separation in planar systems such as benzene. As one can see from the energies of Table II, for a given value of N_{det}, the variational energy and the PT2-corrected energies are much lower with localized orbitals than with natural orbitals. Indeed, localized orbitals significantly speed up the convergence of SCI calculations by taking benefit of the local character of electron correlation.[17,66–70] We, therefore, consider these energies more trustworthy, and we will base our best estimate of the correlation energy of benzene on these calculations. The convergence of the CIPSI correlation energy using localized orbitals is illustrated in Fig. 1, where one can see the behavior of the correlation energy, $\Delta E_{var.}$, and $\Delta E_{var.} + E_{(r)PT2}$, as a function of N_{det} (left panel). The right panel of Fig. 1 is more instructive as it shows $\Delta E_{var.}$ as a function of $E_{(r)PT2}$ and their corresponding four-point linear extrapolation curves that we have used to get our final estimate of the correlation energy. (In other words, the four largest variational wave functions are considered to perform the linear extrapolation.) From this figure, one clearly sees that the rPT2-based correction behaves more linearly than its corresponding PT2 version and is, thus, systematically employed in the following.

Our final number is gathered in Table III, where, following the notations of Ref. [17], we report, in addition to the final variational energies $\Delta E_{var.}$, the extrapolation distances, ΔE_{dist}, defined

TABLE II. Variational energy $E_{var.}$, second-order perturbative correction E_{PT2}, and its renormalized version E_{rPT2} (in E_h) as a function of the number of determinants N_{det} for the ground state of the benzene molecule computed in the cc-pVDZ basis set. The statistical error on $E_{(r)PT2}$, corresponding to one standard deviation, is reported in parentheses.

N_{det}	Natural orbitals			Localized orbitals		
	$E_{var.}$	$E_{var.} + E_{PT2}$	$E_{var.} + E_{rPT2}$	$E_{var.}$	$E_{var.} + E_{PT2}$	$E_{var.} + E_{rPT2}$
1280	−230.978 056	−231.559 025(212)	−231.463 633(177)	−231.101 676	−231.519 522(149)	−231.472 224(132)
2560	−231.043 712	−231.542 344(139)	−231.474 885(120)	−231.161 264	−231.515 577(155)	−231.482 477(140)
5120	−231.115 142	−231.534 122(213)	−231.488 815(190)	−231.224 632	−231.516 375(191)	−231.495 022(177)
10 240	−231.188 813	−231.531 660(516)	−231.502 992(473)	−231.283 295	−231.520 907(271)	−231.507 708(255)
20 480	−231.260 065	−231.534 172(611)	−231.517 063(573)	−231.330 209	−231.526 433(586)	−231.518 045(561)
40 960	−231.321 906	−231.538 269(501)	−231.528 301(478)	−231.366 008	−231.532 288(303)	−231.526 639(293)
81 920	−231.366 895	−231.541 945(813)	−231.535 785(785)	−231.392 888	−231.536 578(614)	−231.532 575(597)
163 840	−231.392 866	−231.545 499(761)	−231.541 010(739)	−231.414 132	−231.541 400(624)	−231.538 378(609)
327 680	−231.407 802	−231.548 699(662)	−231.544 980(645)	−231.431 952	−231.545 873(557)	−231.543 532(545)
655 360	−231.418 752	−231.551 208(661)	−231.548 004(645)	−231.447 007	−231.548 856(498)	−231.547 043(489)
1 310 720	−231.428 852	−231.552 760(616)	−231.550 006(603)	−231.460 970	−231.552 137(453)	−231.550 723(446)
2 621 440	−231.439 324	−231.553 845(572)	−231.551 544(560)	−231.473 751	−231.555 261(403)	−231.554 159(397)
5 242 880	−231.450 156	−231.557 541(534)	−231.555 558(524)	−231.485 829	−231.558 303(362)	−231.557 451(358)
10 485 760	−231.461 927	−231.559 390(481)	−231.557 796(474)	−231.497 515	−231.562 568(322)	−231.561 901(319)
20 971 520	−231.474 019	−231.561 315(430)	−231.560 063(424)	−231.508 714	−231.564 707(275)	−231.564 223(273)
41 943 040	−231.487 978	−231.564 529(382)	−231.563 593(377)	−231.519 122	−231.567 419(240)	−231.567 069(238)
83 886 080	−231.501 334	−231.566 994(317)	−231.566 325(314)	−231.528 568	−231.570 084(199)	−231.569 832(198)
167 772 160	−231.514 009	−231.569 939(273)	−231.569 467(271)	−231.536 655	−231.571 981(175)	−231.571 804(174)

as the difference between the final computed energy, ΔE_{final}, and the extrapolated energy, $\Delta E_{extrap.}$, associated with ASCI, iCI, SHCI, DMRS, and CIPSI. The three flavors of SCI fall into an interval ranging from −860.0 mE_h (ASCI) to −864.2 mE_h (SHCI), while the other non-SCI methods yield correlation energies ranging from −863.7 to

−862.8 mE_h (see Table I). Our final CIPSI number (obtained with localized orbitals and rPT2 correction via a four-point linear extrapolation) is −863.4(5) mE_h, where the error reported in parentheses represents the fitting error (not the extrapolation error for which it is much harder to provide a theoretically sound estimate).[71] For
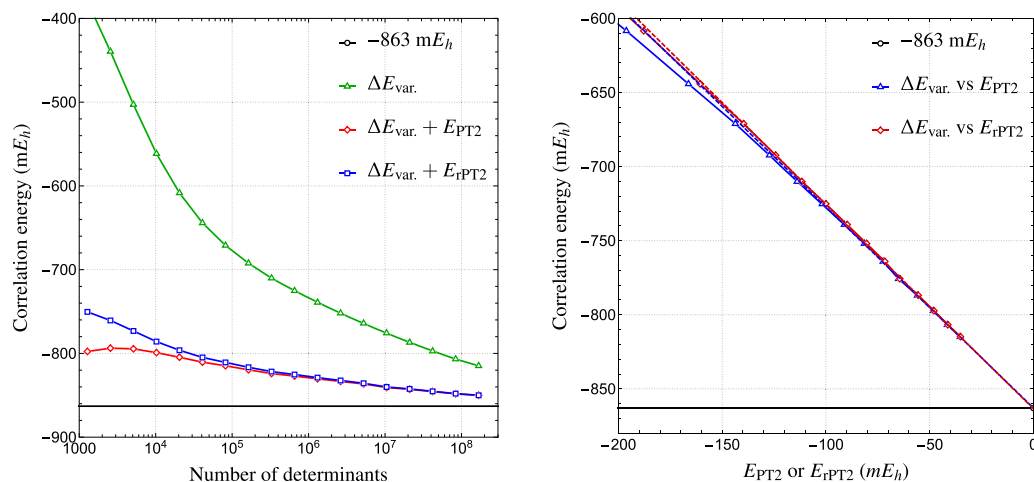

FIG. 1. Convergence of the CIPSI correlation energy of benzene using localized orbitals. Left: $\Delta E_{var.}$, $\Delta E_{var.} + E_{PT2}$, and $\Delta E_{var.} + E_{rPT2}$ (in mE_h) as functions of the number of determinants in the variational space N_{det}. Right: $\Delta E_{var.}$ (in mE_h) as a function of E_{PT2} or E_{rPT2}. The four-point linear extrapolation curves (dashed lines) are also reported. The theoretical estimate of −863 mE_h from Ref. 17 is marked by a black line for comparison purposes. The statistical error bars associated with E_{PT2} or E_{rPT2} (not shown) are of the order of the size of the markers.

TABLE III. Extrapolation distances, ΔE_{dist}, defined as the difference between the final computed energy, $\Delta E_{\mathrm{final}}$, and the extrapolated energy, $\Delta E_{\mathrm{extrap.}}$, associated with ASCI, iCI, SHCI, DMRG, and CIPSI for the best blind-test and post-blind-test estimates of the correlation energy of benzene in the cc-pVDZ basis. The final variational energies $\Delta E_{\mathrm{var.}}$ are also reported. See Ref. 17 for more details. All correlation energies are given in mE_h.

Method	$\Delta E_{\mathrm{var.}}$	$\Delta E_{\mathrm{final}}$	$\Delta E_{\mathrm{extrap.}}$	ΔE_{dist}
Best blind-test estimates				
ASCI	−737.1	−835.4	−860.0	−24.6
iCI	−730.0	−833.7	−861.1	−27.4
SHCI	−827.2	−852.8	−864.2	−11.4
DMRG	−859.2	−859.2	−862.8	−3.6
Best post-blind-test estimates				
ASCI	−772.4	−835.2	−861.3	−26.1
iCI	−770.7	−842.8	−864.2	−21.3
SHCI	−835.2	−854.9	−863.6	−8.7
CIPSI	−814.8	−850.2	−863.4	−13.2

comparison, the best post-blind-test SHCI estimate is −863.3 mE_h, which agrees almost perfectly with our best CIPSI estimate, while the best post-blind-test ASCI and iCI correlation energies are −861.3 and −864.15 mE_h, respectively (see Table III).

The present calculations have been performed on the AMD partition of GENCI's Irene supercomputer. Each Irene's AMD node is a dual-socket AMD Rome (EPYC) CPU at 2.60 GHz with 256GiB of RAM, with a total of 64 physical cores per socket. These nodes are connected via Infiniband HDR100. The first step of the calculation, i.e., performing a CIPSI calculation up to $N_{\mathrm{det}} \sim 10^7$ with Hartree–Fock orbitals in order to produce natural orbitals, takes roughly 24 h on a single node, and reaching the same number of determinants with natural orbitals or localized orbitals takes roughly the same amount of time. A second 24-h run on 10 distributed nodes was performed to push the selection to 8×10^7 determinants, and a third distributed run using 40 nodes was used to reach 16×10^7 determinants. In total, the present calculation has required 150 k core hours, most of it being spent in the last stage of the computation.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Zenodo at http://doi.org/10.5281/zenodo.4075286.

REFERENCES

[1] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos et al., Phys. Rev. X 5, 041041 (2015).

[2] M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. A. Jiménez-Hoyos, T. N. Lan, J. Li, F. Ma et al., Phys. Rev. X 7, 031059 (2017).

[3] K. T. Williams, Y. Yao, J. Li, L. Chen, H. Shi, M. Motta, C. Niu, U. Ray, S. Guo, R. J. Anderson et al., Phys. Rev. X 10, 011041 (2020).

[4] S. P. A. Sauer, M. Schreiber, M. R. Silva-Junior, and W. Thiel, J. Chem. Theory Comput. 5, 555 (2009).

[5] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. 128, 134110 (2008).

[6] M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. 129, 104103 (2008).

[7] M. R. Silva-Junior, S. P. A. Sauer, M. Schreiber, and W. Thiel, Mol. Phys. 108, 453 (2010).

[8] M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. 133, 174318 (2010).

[9] C. F. Bender and E. R. Davidson, Phys. Rev. 183, 23 (1969).

[10] B. Huron, J. P. Malrieu, and P. Rancurel, J. Chem. Phys. 58, 5745 (1973).

[11] R. J. Buenker and S. D. Peyerimhoff, Theor. Chim. Acta 35, 33 (1974).

[12] P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin, J. Chem. Theory Comput. 14, 4360 (2018).

[13] P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, and D. Jacquemin, J. Chem. Theory Comput. 15, 1939 (2019).

[14] P.-F. Loos, A. Scemama, and D. Jacquemin, J. Phys. Chem. Lett. 11, 2374 (2020).

[15] P.-F. Loos, F. Lipparini, M. Boggio-Pasqua, A. Scemama, and D. Jacquemin, J. Chem. Theory Comput. 16, 1711 (2020).

[16] P.-F. Loos, A. Scemama, M. Boggio-Pasqua, and D. Jacquemin, J. Chem. Theory Comput. 16, 3720 (2020).

[17] J. J. Eriksen, T. A. Anderson, J. E. Deustua, K. Ghanem, D. Hait, M. R. Hoffmann, S. Lee, D. S. Levine, I. Magoulas, J. Shen, N. M. Tubman, K. B. Whaley, E. Xu, Y. Yao, N. Zhang, A. Alavi, G. K.-L. Chan, M. Head-Gordon, W. Liu, P. Piecuch, S. Sharma, S. L. Ten-no, C. J. Umrigar, and J. Gauss, J. Phys. Chem. Lett. 11, 8922 (2020).

[18] N. Oliphant and L. Adamowicz, J. Chem. Phys. 95, 6645 (1991).

[19] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. 97, 4282 (1992).

[20] J. J. Eriksen, F. Lipparini, and J. Gauss, J. Phys. Chem. Lett. 8, 4633 (2017).

[21] J. J. Eriksen and J. Gauss, J. Chem. Theory Comput. 14, 5180 (2018).

[22] J. J. Eriksen and J. Gauss, J. Chem. Theory Comput. 15, 4873 (2019).

[23] J. J. Eriksen and J. Gauss, J. Phys. Chem. Lett. 10, 7910 (2019).

[24] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, J. Chem. Phys. 145, 044112 (2016).

[25] N. M. Tubman, D. S. Levine, D. Hait, M. Head-Gordon, and K. B. Whaley, "An efficient deterministic perturbation theory for selected configuration interaction methods," arXiv:1808.02049 [cond-mat.str-el] (2018).

[26] N. M. Tubman, C. D. Freeman, D. S. Levine, D. Hait, M. Head-Gordon, and K. B. Whaley, J. Chem. Theory Comput. 16, 2139 (2020).

[27] W. Liu and M. R. Hoffmann, Theor. Chem. Acc. 133, 1481 (2014).

[28] W. Liu and M. R. Hoffmann, J. Chem. Theory Comput. 12, 1169 (2016).

[29] Y. Lei, W. Liu, and M. R. Hoffmann, Mol. Phys. 115, 2696 (2017).

[30] N. Zhang, W. Liu, and M. R. Hoffmann, J. Chem. Theory Comput. 16, 2296 (2020).

[31] A. A. Holmes, H. J. Changlani, and C. J. Umrigar, J. Chem. Theory Comput. 12, 1561 (2016).

[32] A. A. Holmes, C. J. Umrigar, and S. Sharma, J. Chem. Phys. 147, 164111 (2017).

[33] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, J. Chem. Theory Comput. 13, 1595 (2017).

[34] E. Xu, M. Uejima, and S. L. Ten-no, Phys. Rev. Lett. 121, 113001 (2018).

[35] E. Xu, M. Uejima, and S. L. Ten-no, "Towards near-exact solutions of molecular electronic structure: Full coupled-cluster reduction with a second-order perturbative correction," arXiv:2010.01850 [physics.chem-ph] (2020).

[36]S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[37]S. R. White, Phys. Rev. B **48**, 10345 (1993).

[38]G. K.-L. Chan and S. Sharma, Annu. Rev. Phys. Chem. **62**, 465 (2011).

[39]G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[40]D. Cleland, G. H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[41]K. Ghanem, A. Y. Lozovoi, and A. Alavi, J. Chem. Phys. **151**, 224108 (2019).

[42]J. E. Deustua, J. Shen, and P. Piecuch, Phys. Rev. Lett. **119**, 223003 (2017).

[43]J. E. Deustua, I. Magoulas, J. Shen, and P. Piecuch, J. Chem. Phys. **149**, 151101 (2018).

[44]M. Motta and S. Zhang, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **8**, e1364 (2018).

[45]J. Lee, F. D. Malone, and D. R. Reichman, J. Chem. Phys. **153**, 126101 (2020).

[46]S. Evangelisti, J.-P. Daudey, and J.-P. Malrieu, Chem. Phys. **75**, 91 (1983).

[47]E. Giner, A. Scemama, and M. Caffarel, Can. J. Chem. **91**, 879 (2013).

[48]E. Giner, A. Scemama, and M. Caffarel, J. Chem. Phys. **142**, 044115 (2015).

[49]Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, J. Chem. Phys. **147**, 034101 (2017).

[50]Y. Garniron, A. Scemama, E. Giner, M. Caffarel, and P.-F. Loos, J. Chem. Phys. **149**, 064103 (2018).

[51]Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner, and A. Scemama, J. Chem. Theory Comput. **15**, 3591 (2019).

[52]M. Caffarel, E. Giner, A. Scemama, and A. Ramírez-Solís, J. Chem. Theory Comput. **10**, 5286 (2014).

[53]M. Caffarel, T. Applencourt, E. Giner, and A. Scemama, J. Chem. Phys. **144**, 151103 (2016).

[54]M. Caffarel, T. Applencourt, E. Giner, and A. Scemama, "Using CIPSI nodes in diffusion Monte Carlo," in *Recent Progress in Quantum Monte Carlo*, ACS Symposium Series Vol. 1234 (American Chemical Society, 2016), Chap. 2, pp. 15–46.

[55]A. Scemama, E. Giner, T. Applencourt, and M. Caffarel, "QMC using very large configuration interaction-type expansions," in *Advances in Quantum Monte Carlo* (Pacifichem, 2015).

[56]A. Scemama, T. Applencourt, E. Giner, and M. Caffarel, J. Comput. Chem. **37**, 1866 (2016).

[57]A. Scemama, Y. Garniron, M. Caffarel, and P.-F. Loos, J. Chem. Theory Comput. **14**, 1395 (2018).

[58]A. Scemama, A. Benali, D. Jacquemin, M. Caffarel, and P.-F. Loos, J. Chem. Phys. **149**, 034108 (2018).

[59]A. Scemama, M. Caffarel, A. Benali, D. Jacquemin, and P.-F. Loos, Results Chem. **1**, 100002 (2019).

[60]M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018).

[61]M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **15**, 4896 (2019).

[62]T. Applencourt, K. Gasperich, and A. Scemama, "Spin adaptation with determinant-based selected configuration interaction," arXiv:1812.06902 [physics.chem-ph] (2018).

[63]J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).

[64]Indices of molecular orbitals for the Boys–Foster localization procedure: core [1–6]; σ [7–18]; π [19–21]; π^* [22–24]; σ^* [25–36]; higher-lying π [39, 41–43, 46, 49, 50, 53–57, 71–74, 82–85, 87, 92, 93, 98]; and higher-lying σ [37, 38, 40, 44, 45, 47, 48, 51, 52, 58–70, 75–81, 86, 88–91, 94–97, 99–114].

[65]J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[66]C. Angeli, C. J. Calzado, R. Cimiraglia, S. Evangelisti, N. Guihéry, T. Leininger, J.-P. Malrieu, D. Maynau, J. V. P. Ruiz, and M. Sparta, Mol. Phys. **101**, 1389 (2003).

[67]C. Angeli, J. Comput. Chem. **30**, 1319 (2009).

[68]N. Ben Amor, F. Bessac, S. Hoyau, and D. Maynau, J. Chem. Phys. **135**, 014101 (2011).

[69]N. Suaud and J.-P. Malrieu, Mol. Phys. **115**, 2684 (2017).

[70]A. D. Chien, A. A. Holmes, M. Otten, C. J. Umrigar, S. Sharma, and P. M. Zimmerman, J. Phys. Chem. A **122**, 2714 (2018).

[71]Using the 3, 4, 5, and 6 largest wave functions to perform the linear extrapolation yields the following correlation energy estimates: $-863.1(11)$, $-863.4(5)$, $-862.1(8)$, and $-863.5(11)$ mE_h, respectively. These numbers vary by 1.4 mE_h. The four-point extrapolated value of $-863.4(5)$ mE_h that we have chosen to report as our best estimate corresponds to the smallest fitting error. Quadratic fits yield much larger variations and are discarded in practice. Due to the stochastic nature of E_{rPT2}, the fifth point is slightly off as compared to the others. Taking into account this fifth point yields a slightly smaller estimate of the correlation energy [$-862.1(8)$ mE_h], while adding a sixth point settles down the correlation energy estimate at $-863.5(11)$ mE_h.