

Computer-predicted excitation energy of carbon in agreement with experiment to within 1 cm^{-1}

Nike Dattani*

*Oxford University, Hertford College, OX1 3BW, Oxford, UK, and
Kyoto University, Fukui Institute for Fundamental Chemistry, 606-8103, Kyoto, Japan,*

David Feller†

Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA,

Jacek Koput‡

Department of Chemistry, Adam Mickiewicz University, 60-780 Poznan, Poland.

and

(Dated: 8th May 2017)

We show that with non-explicitly correlated basis sets, we can predict the first excitation energy in carbon to within 1 cm^{-1} of the experimental value.

I. INTRODUCTION

In the last three years, excitation energies have been calculated with unprecedented precision for the Li atom [?], Be atom [?] and B atom [?]. Tight variational bounds for non-relativistic ground state energies assuming a clamped, point-sized nucleus have reached 49 digits in units of Hartree for He, 19 digits for Li, 12 digits for Be, 11 digits for B, and 6 digits for C (see Table I). Calculated excitation energies have been made in agreement with experiment to within 10^{-3} cm^{-1} for ^7Li , 10^{-1} cm^{-1} for ^9Be and 1 cm^{-1} for ^{10}B (see Table II).

For the C atom, before this present study, no high-precision calculation had been reported to predict an excitation energy to within 1 cm^{-1} agreement with experiment. Table I shows that the method used for the calculation has a crucial impact on the precision obtainable: the best variational bound on the non-relativistic, clamped, point-nucleus (NR,CPN), ground state energy of Li is believed to be 8 orders of magnitude more accurate than that of the Be atom, because analytic expressions for the multi-center integrals of Hylleraas wavefunctions are only known for up to $3e^-$, and it would be very slow to calculate them numerically for $4e^-$. For $4e^-$ and greater, ECGs (explicitly correlated gaussian wavefunctions) are used. ECGs mimic the theoretical shape of the wavefunction less accurately (e.g. they are rounded rather than cusp-like in the limit of vanishing electron-electron distance, and in the infinite electron-electron distance limit they decay as $1/r^2$ rather than $1/r$), so more parameters are needed than when using Hylleraas wavefunctions. However, the integrals of ECGs are known analytically, regardless of the number of electrons, which is why it has been possible to do very high-precision calculations on Be [?] and B [?].

Table I demonstrates though, that with twice as many variationally optimizable parameters, the authors obtained convergence in 1 fewer digit for the B atom than for the Be atom. This suggests that it would be very difficult to use the ECG approach for atoms and molecules coming from most of the periodic table. The best known variational bound for the NR,CPN ground state energy for C was calculated in 2015 using fixed-node diffusion Monte Carlo (FN-DMC) with the nodes of the electronic wavefunction fixed at the locations of the CISD/cc-pV5Z wavefunction, and the statistical uncertainty on the Monte Carlo was $\pm 2 E_h$. However, FN-DMC has not yet been able to calculate the energies of excited states. In this paper, the approach we use to calculate the NR,CPN energy of the ground state of C is FCIQMC (full configuration interaction Monte Carlo) in an aug-cc-pCV7Z basis set. Fig I shows that our NR,CPN energy is at least $210\text{ }\mu E_h$ higher than the variational upper bound obtained from FN-DMC; but since it is straightforward to calculate excited state energies with FCIQMC, unlike FN-DMC, we are easily able to calculate the first excitation energy for carbon ($2^2P \leftarrow 2^1D$) with high precision in this work.

After adding estimates for the correction to the non-relativistic approximation, and for the correction to the clamped nucleus approximation, we obtained a ($2^2P \leftarrow 2^1D$) excitation energy for C which is within 1 cm^{-1} disagreement with the best known experimental estimate for this transition. While this is not as impressive as the method of variationally optimizing parameters in an explicitly correlated wavefunction ansatz has proven to be for Li and Be, the disagreement has the same order of magnitude as the latter approach for B (see Table II). We finally note that the approach used in this paper, of calculating FCIQMC on a basis set of non-explicitly correlated orbitals has successfully treated systems with far more electrons (transition metal atoms [?], diatomics [?], larger molecules such as butadiene [?], and even solid state systems [?]), so it is conceivable that the approach used in this paper may in the near future be able

* dattani.nike@hertford.ox.ac.uk

† dfeller@owt.com

‡ koput@amu.edu.pl

to determine (with good accuracy) the first excitation energy for astatine, which at present remains experimentally elusive.

II. METHODOLOGY

Our calculations in this paper can be divided into four stages: (A) We develop more accurate basis sets than previously available for carbon, (B) we calculate the 1- and 2-electron integrals in these basis sets, (C) we attempt to solve the NR-CPN Schroedinger equation at the FCI level, and (D) we estimate the size of corrections due to special relativity and due to the atom having an unclamed nucleus.

A. Optimization of ‘tight function’ exponents for the aug-cc-pCV7Z and aug-cc-pCV8Z basis sets.

The largest basis sets known for C prior to this work were the (aug-cc-pVXZ, $X=7,8,9$) sets used by Feller earlier this year. These basis sets did not contain ‘tight’ exponent functions for capturing the effects of the correlation between the core ($1s^2, 2s^2$) electrons and the valence electrons ($2p^2$). The largest known basis set prior to this work including the CV (core-valence) correction was the aug-cc-pCV6Z set. In this work we start by optimizing the ‘tight’ exponents for the CV correction to Feller’s 2016 aug-cc-pV7Z and aug-cc-pV8Z basis sets, yielding the first aug-cc-pCV7Z basis set for carbon, and the first aug-cc-pCV8Z basis set known.

Optimizing exponents for such large basis sets can be quite a challenging task. We employ two different approaches for the optimization of the ‘tight’ functions in our CV corrections. [David’s approach, and Jacek’s approach].

B. Calculation of 1- and 2-electron integrals including k - and l - functions

The calculation of the 1- and 2-electron integrals for (aug)-cc-p(C)VXZ basis sets with $X \geq 7$ is not possible with most quantum chemistry packages, since very few software packages can write 1- and 2-electron integrals in FCIDUMP format, while also supporting k - and l - functions, but k -functions appear in $X = 7$ basis sets and l -functions appear when $X = 8$. We have used MOLCAS to write the FCIDUMP files, and we used a CAS(6,6) for the 1D state.... The commercial version of MOLCAS had

a memory leak when printing the 56 GB FCIDUMP file for aug-cc-pCV8Z, so we had to make some changes to MOLCAS ...

C. Calculation of non-relativistic energies including all possible levels of excitation (with FCIQMC)

We expand the wavefunction of the relevant electronic state as a sum of all possible Slater determinants (full configuration interaction), and the coefficients are determined by the number of walkers standing on each determinant after a Monte Carlo sampling using the NR-CPN Hamiltonian. The method was introduced in [?], and we use the initiator method first described in [?], and the semi-stochastic method as described in [?]. The calculations are performed using Version X.X of the software NECI.

Within a given Hamiltonian (in this case the NR-CPN Hamiltonian) and basis set, there are three sources of error in the FCIQMC energy calculations:

1. Trial wavefunction error (ΔE_{trial}), which approaches zero in the limit where the number of determinants used in the trial wavefunction approaches the number of determinants in the FCIQMC wavefunction;
2. Initiator error ($\Delta E_{\text{initiator}}$), which approaches zero in the limit where the number of walkers N_{walkers} gets sufficiently large; and
3. Stochastic error (ΔE_{stoch}), which for a given number of walkers is estimated as the square root of the unbiased variance among different estimates E_i of the energy from their mean \bar{E} after different numbers N of Monte Carlo iterations after the walkers have reached equilibrium: $\Delta E_{\text{stochastic}} \approx \sqrt{\frac{\sum_{i=1}^N (E_i - \bar{E})^2}{N-1}} = \mathcal{O}(1/\sqrt{N})$.

Our goal was to obtain all energies to a precision of $\pm\epsilon$ where $\epsilon \leq 1\mu E_{\text{Hartree}} \approx 0.2 \text{ cm}^{-1}$ (within the aug-cc-pCV7Z basis set used). Therefore, every calculation was run for enough iterations N such that ΔE_{stoch} was smaller than $1\mu E_{\text{Hartree}}$. To ensure that $\Delta E_{\text{initiator}}$ can be neglected, we used a sufficiently large value of N_{walkers} for every energy calculation, so that the energy difference between using N_{walkers} and $\frac{1}{2}N_{\text{walkers}}$ was at least an order of magnitude smaller than $\Delta E_{\text{stochastic}}$. Likewise, to ensure that $\Delta E_{\text{estimator}}$ can be neglected, we used a sufficiently large number of determinants for every energy calculation, such that $\Delta E_{\text{estimator}}$ would be at least an order of magnitude smaller than $\Delta E_{\text{stochastic}}$.

D. Estimation of relativistic corrections and finite nuclear mass corrections

See other PDF.

Table I. Tight variational upper bounds for atomic energies. For H, the listed energy is exact, since the Hartree energy unit is defined as twice this energy, and can be related to fundamental SI units through an analytic solution to the relevant Schrödinger equation. VO stands for variational optimization (parameters in a wavefunction ansatz are optimized in attempt to obtain the lowest possible energy). Hylleraas-Log indicates the use of Hylleraas functions supplemented with auxiliary log functions, and ECG(N) indicates the use of explicitly correlated Gaussians with N variationally optimizable parameters. FN-DMC stands for fixed node diffusion Monte Carlo.

Total non-relativistic clamped point-nucleus (NR,CPN) energy [Hartree]										Method/Ansatz type	Reference
H	1	-0.5								Analytic	1926 Schroedinger
He	2	-2.903 724 377 034 119 598 311 159 245 194 404 446 696 925 309 838								VO/Hylleraas-Log	2006 Schwartz
Li	3	-7.478 060 323 910 134 843								VO/Hylleraas	2011 Wang
Be	4	-14.667 356 494 9								VO/ECG(4096)	2013 Puchalski [?]
B	5	-24.653 867 537								VO/ECG(8192)	2015 Puchalski [?]
C	6	-37.844 48(2)								FN-DMC	2015 Yang
C	6	-37.844 251 5(08)								FCIQMC/aCV7Z	Present work -
C	6	-37.843 333								VO/ECG(1000)	2013 Bubin [?]

Table II. The most precisely calculated electronic excitation energies for the first 6 atoms, compared to the best known experimental measurements to date. The last column indicates that if aiming for the best precision, an experimental measurement is still the best way to obtain the energy for most atoms, but for Be, the energy has been obtained more precisely *in silico* than in any experiment to date.

		Experiment	Theory	Calc - Obs	$ \frac{\text{Calc} - \text{Obs}}{\text{Uncertainty in obs}} $	More precise	
		[cm ⁻¹]	[cm ⁻¹]	[cm ⁻¹]			
<hr/>							
<hr/>							
H							
He							
³ Li	$2^2P_{3/2} \leftarrow 2^2S_{1/2}$	14 903.632 061 017 4(5) [?]	14903.631 765(7)	[?]	-0.000 296	592	Experiment
⁹ Be	$2^1P \leftarrow 2^1S$	42 565.35(18)	42 565.441(11)	[?]	0.091	0.51	Theory
¹¹ B	$3^2P_{1/2} \leftarrow 2^2P_{1/2}$	40 039.656(3)	40 039.52(24)	[?]	-0.136	45	Experiment
C	$2^3P \leftarrow 2^1D$	10 192.654(100)	10 191.6(9)	Present work			Experiment
<hr/>							

Table III. Basis set and correlation convergence of the X2C and DBOC corrections to the $\text{C}(^3P) \rightarrow \text{C}^+(^2P)$ ionization energy in cm⁻¹. The difference between CCSDT and FCI for the smallest two basis sets is denoted by $\Delta_{\text{correlation}}$ and the values at aCV3Z are added to the CCSDT energies for all larger basis sets (resulting in FCI estimates presented in *italic* font). The numbers in parentheses for the CBS (complete basis set) estimates denote our assigned uncertainties in the last digits presented, which we believe to be very conservative. aCVXZ is short for aug-cc-pCVXZ-uncontracted.

	aCV2Z	aCV3Z	aCV4Z	aCV5Z	aCV6Z	CBS
X2C						
CCSDT	-31.055	-30.144	-30.037	-30.004	-30.001	-29.999(003)
FCI	-31.074	-30.168	<i>-30.060</i>	<i>-30.028</i>	<i>-30.025</i>	<i>-30.023(050)</i>
$\Delta_{\text{correlation}}$	-0.019	-0.024	<i>-0.024</i>	<i>-0.024</i>	<i>-0.024</i>	<i>-0.024(050)</i>
DBOC						
CCSDT	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
FCI	-0.2	-0.2	<i>-0.2</i>	<i>-0.2</i>	<i>-0.2</i>	<i>-0.2</i>
$\Delta_{\text{correlation}}$	0.0	0.0	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>

III. RESULTS

IV. DISCUSSION

Table IV. Tight exponents for aug-cc-pCV6Z and aug-cc-pCV7Z.

	aug-cc-pCV7Z	aug-cc-pCV8Z
<i>s</i> -type	276.12 158.30 90.75 52.026 29.826 17.099	365.52 232.20 147.51 93.707 59.529 37.817 24.024
<i>p</i> -type	299.20 149.46 74.657 37.293 18.629 9.3054	372.10 198.95 106.37 56.868 30.405 16.256 8.6911
<i>d</i> -type	255.63 111.57 55.17 27.282 13.491	337.98 191.20 108.17 61.193 34.618 19.584
<i>f</i> -type	132.26 56.375 24.030 10.243	207.91 103.16 51.185 25.397 12.601
<i>g</i> -type	94.488 36.762 14.302	155.52 72.601 33.891 15.821
<i>h</i> -type	66.227 22.68	100.23 39.552 15.608
<i>i</i> -type	48.32	78.691 28.941
<i>k</i> -type		

Table V. FCIQMC

			$E = \frac{\langle \psi_{N_{\text{trial}}} H_{\text{NR-CPN}} \psi_{\text{FCIQMC}} \rangle}{\langle \psi_{N_{\text{trial}}} \psi_{\text{FCIQMC}} \rangle}$	$E_{\text{excitation}}$	
N_{walkers}	N_{trial}	Uncertainty	2^3P [E_{hartree}]	2^1D [E_{hartree}]	$2^3P \rightarrow 2^1D$ [cm^{-1}]
aug-cc-pCV6Z					
64×10^6	1		-37.843 839 6(18)		
	1000		-37.843 840 5(05)		
128×10^6	1	ΔE_{trial}	< 0.000 001 1		
	1000		-37.843 839 2(99)	-37.	
			-37.843 840 5(05)		
		ΔE_{trial}	< 0.000 001 3		
		$\Delta E_{\text{initiator}}$	< 0.000 000 4		
aug-cc-pCV7Z					
64×10^6	1		-37.844 251 5(15)	-37.430 345 0(03)	90 841.955
	1000		-37.844 251 5(08)	-37.430 345 0(01)	90 841.977
		ΔE_{trial}	< 0.000 000 1	< 0.000 000 1	< 0.02
128×10^6	1		-37.844 251 5(0?)	-37.430 345 0(02)	90 841.977
	1000		-37.844 251 5(05)	-37.430 345 1(01)	90841.955
		ΔE_{trial}	< 0.000 000 1	< 0.000 000 1	< 0.02
		$\Delta E_{\text{initiator}}$	< 0.000 000 1	< 0.000 000 1	< 0.02
CBS Extrapolation					
$1/X^3$			-37.844 697 1	-37.430 741 9	90 852.665
$1/X^{3.5}$			-37.844 645 6	-37.430 696 0	90 851.416
$1/(X + 1/2)^4$			-37.844 625 1	-37.430 677 8	90 850.925
		$\Delta_{\text{basis set}}$	\sim 0.000 037 0	\sim 0.000 033 0	0.90
Non-relativistic, clamped nucleus (FCIQMC)					90 851.69(90)
X2C Relativistic correction					-30.02(05)
Breit + QED (from)					-0.48
DBOC unclamped nucleus correction					0.2
Theory	(Present Work)	FCIQMC + X2C + Breit + QED + DBOC			90 820.76(95)
Experiment	(1990)				90 820.45(10)
Theory	(2010)	CCSD-F12 + HLC + MV + Darwin + Breit + QED + DBOC			90 826.09
Theory	(2017)				90 827.50
Theory	(2015)	FN-DMC			90 774.