

Excited States of Methylene, Polyenes, and Ozone from Heat-bath Configuration Interaction

Alan D. Chien,¹ Adam A. Holmes,²,³ Matthew Otten,³ C. J. Umrigar,³ Sandeep Sharma,² and Paul M. Zimmerman⁴

¹Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA
²Department of Chemistry and Biochemistry, University of Colorado Boulder, Boulder, CO 80302, USA
³Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY 14853, USA
⁴Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

The electronically excited states of methylene (CH2), ethylene (C2H4), butadiene (C4H6), hexatriene (C6H3), and ozone (O3) have long proven challenging due to their complex mixtures of static and dynamic correlations. Semistochastic heat-bath configuration interaction (SHCI), which efficiently and systematically approaches the full configuration interaction (FCI) limit, is used to provide close approximations to the FCI energies in these systems. This article presents the largest FCI-level calculation to date – on hexatriene using a polarized double-zeta basis (ANO-L-pVDZ), which gives rise to a Hilbert space containing more than 10¹⁶ determinants. These calculations give vertical excitation energies of 5.58 and 5.59 eV respectively for the 2¹A₂ and 1¹B₃ states, showing that they are nearly degenerate. The same excitation energies in butadiene/ANO-L-pVDZ were found to be 6.58 and 6.45 eV. In addition to these benchmarks, our calculations strongly support the presence of a previously hypothesized ring-minimum species of ozone that lies 1.3 eV higher than the open-ring minimum energy structure and is separated from it by a barrier of 1.11 eV.

I. INTRODUCTION

The exponential increase in Hamiltonian dimension with increasing system size means that exact, Born-Oppenheimer electronic energies are not easily achievable for polyatomic molecular systems.¹² Recently, several methods that produce FCI-quality energies at greatly reduced cost, such as density matrix renormalization group (DMRG)⁵⁻⁷, FCI quantum Monte Carlo (FCIQMC),¹³⁻¹⁵ and incremental FCI (IFCI)¹⁷⁻¹⁹

Another avenue for obtaining FCI-quality energies has also recently become available due to the revival of the semistochastic heat-bath configuration interaction (SHCI), which has been shown to efficiently treat CI spaces orders of magnitude larger than is possible with conventional CI algorithms.⁵²⁻⁵⁵

In this paper, highly accurate benchmarks for electronically excited states of the polyatomics in Figure 1 are computed using SHCI. Methylene is presented as the first test case, due to its small size yet challenging electronic structure.⁵⁶ Additionally, ozone is examined to answer a long-standing question regarding the existence of a theoretically meta-stable ozone species.⁵⁷⁻⁵⁸ Finally, SHCI is applied to the first few polyenes – ethylene, butadiene, and hexatriene – which have long been studied for their role as prototypical organic conducting polymers. The state ordering of the low-lying valence excited states, 2¹A₂ and 1¹B₃, in butadiene and hexatriene have been especially challenging due to the significant numbers of highly correlated electrons and the near-degenerate nature of the valence states.⁵⁹⁻⁶¹ Herein, extrapolated SHCI state energies will provide high accuracy 2¹A₂/1¹B₃ state orderings in butadiene and hexatriene.

This article is organized as follows. In Section II, the excited-state SHCI algorithm is reviewed, and the path to convergence to the FCI limit is discussed. In Section III, results on the smaller methylene and ethylene systems are used to validate SHCI against current benchmark values and establish convergence with respect to FCI. These observations are then used to estimate FCI-quality energies for the larger ozone, butadiene, and hex-
have been described in detail,\textsuperscript{51–53} to the final perturbation only considers the determinants connected through threshold values, $\epsilon_V$ and $\epsilon_{PT}$, respectively. The current variational space of determinants is denoted by $V$ and the space of all determinants connected by single or double excitations to $V$, but not in $V$, is denoted by $C$.

The variational stage iteratively adds determinants to $V$ by

1. Adding all determinants $a$ connected to determinants in the current $V$ that pass the importance criterion $\max_n |H_{ai} \max_i |\langle \psi_n^a |\rangle| > \epsilon_V$, where $c_i^a$ is the coefficient of determinant $i$ in state $n$.

2. Constructing the Hamiltonian and solving for the roots of interest, in the basis of all determinants in the newly expanded $V$.

3. Repeat 1-2 until convergence.

Convergence of the variational wavefunction for a given $\epsilon_V$ is signified by the addition of a small number of new determinants or small changes in the variational energy ($E_{\text{var}}$). The second-order Epstein-Nesbet perturbative energy correction ($\Delta E_2$) is added to $E_{\text{var}}$ to obtain the total HCI energy ($E_{\text{tot}}$). This correction is

$$\Delta E_2 = \sum_{a \in C} \left( \frac{\sum_{i \in V} (\epsilon_{PT}) H_{ai} c_i}{E_0 - H_{aa}} \right)^2, \tag{1}$$

where $a$ runs over determinants in $C$, and $i$ over determinants in $V$. Similar to the variational stage, the perturbation only considers the determinants connected to the final $V$ space that have an importance measure greater than a parameter $\epsilon_{PT}$, which is typically orders of magnitude smaller than $\epsilon_V$. In both the variational and the perturbative stages, the fact that the number of distinct values of the double-excitation matrix elements scales only as $N_{\text{orb}}^4$ is used to avoid ever looking at the unimportant determinants. Nevertheless, storing the full space of determinants used in the perturbative correction becomes a memory bottleneck for larger systems.

SHCI sidesteps this memory bottleneck using a semistochastic second-order perturbation correction.\textsuperscript{50} In this procedure, the perturbative correction is split into deterministic and stochastic contributions. A larger $\epsilon_{PT}$, automatically determined to correspond to a determinant space of manageable size depending on available computer memory, is first used to obtain a deterministic energy correction. The remaining correlation is then calculated stochastically by taking the difference of the second-order corrections evaluated with $\epsilon_{PT}$ and $\epsilon_{PT}^d$. Samples are taken until the statistical error falls below a specified threshold.

### B. Converging SHCI Energies to the FCI Limit

The target accuracy for total or relative energies are typically be chosen to be 1 mHa or 1.6 mHa (1 kcal/mol, representing chemical accuracy), though for the smaller systems it is easy to achieve much higher accuracy. In SHCI, the error in the variational energy can be straightforwardly estimated by the magnitude of the perturbative correction.

In methylene and ethylene, SHCI can provide such highly converged variational energies. In larger systems, however, converging the variational energy would require prohibitively large variational spaces. Instead, we fit the variational energy or the total energy, $E_{\text{tot}} = E_{\text{var}} + \Delta E_2$, to $E_{\text{var}} - E_{\text{tot}}$ using a quadratic function and use the fitted function to extrapolate to the no perturbative correction ($E_{\text{var}} - E_{\text{tot}} = 0$) limit. The fit coefficients for variational and total energies are the same, except that the coefficients of the linear terms differ by one, so the two energies extrapolate to precisely the same value. There is not a well-defined method for estimating the extrapolation error, but a reasonable choice is one fifth of the difference between the calculated energy with the smallest value of $\epsilon_V$ and the extrapolated energy. In many cases the fitted function is very nearly linear (see e.g. Figs. 2, 4 and 5). Furthermore, even when the fitted functions are not close to linear, the functions for different states are often close to parallel, making the estimates of the energy differences particularly accurate.

### C. Computational Details

SHCI is implemented in Fortran90, parallelized using MPI, and makes use of spatial symmetry, and time-reversal symmetry when the number of up- and down-
spin electrons is equal. The variational iterations are terminated when the number of new determinants added is less than 0.001% of the current variational space or when the change in variational energy is less than $1 \cdot 10^{-5}$ Ha. For all calculations, $\epsilon_T$ is set to $1 \cdot 10^{-7}$ Ha, which provides converged perturbative corrections. $\epsilon_V$ is made as small as possible on our hardware, obtaining either small $\Delta E_2$ or enough data points to reliably extrapolate to $\Delta E_2 = 0$. The threshold for statistical error of the stochastic perturbative correction is generally set to $5 \cdot 10^{-5}$ Ha, although the larger hexatriene/ANO-L-pVDZ computations use $1 \cdot 10^{-4}$ Ha.

For the smaller systems (methylene and ethylene) achieving convergence is relatively easy, allowing the use of Hartree-Fock and HCl natural orbitals (obtained with $\epsilon_V = 3 \cdot 10^{-5}$ Ha), for methylene and ethylene respectively, to construct the molecular orbital integrals. For the larger systems (ozone, butadiene, hexatriene), the convergence was improved by using orbitals that minimize the HCl variational energy for $\epsilon_V = 2 \cdot 10^{-4}$. Possibly, yet better convergence could be obtained by using orbitals that make the total energy stationary.

Basis sets used are aug-cc-pVQZ for methylene, ANO-L-pVTZ for ethylene, ANO-L-pVDZ for butadiene and hexatriene, and cc-pVTZ for ozone. Geometries for methylene are FCI/TZVP quality, obtained from Sherrill et al. and ozone geometries are CASSCF(18,12)/cc-pVQZ quality, taken from Theis et al. For the polyenes, all geometries are of MP2/cc-pVQZ quality, with ethylene and hexatriene geometries the same as in Zimmerman and butadiene the same as in Alavi et al. All calculations utilize the frozen-core approximation. For comparisons to coupled cluster theories, the same geometries and basis sets are used with the Q-Chem 4.0 CR-EOM-CC(2,3)D implementation.

### III. RESULTS AND DISCUSSION

#### A. Methylene

Methylene is a prototypical test case for advanced electronic structure methods, being small enough to be amenable to canonical FCI benchmarks, yet still requiring accurate treatment of dynamic and static correlations for correct excitation energies. The four lowest lying states of methylene vary in spin and spatial symmetry: $1^3B_1$, $1^1A_1$, $1^1B_1$, and $2^1A_1$. With only six valence electrons to correlate, SHCI can handle such problems and obtain FCI-quality energies even with the large aug-cc-pVQZ basis (Table 1), obtaining perturbative corrections less than 0.05 mHa with $\epsilon_V = 10^{-5}$ Ha.

Table 1 shows the most accurate SHCI adiabatic energy gaps calculated with $\epsilon_V = 10^{-5}$ Ha, which differ from experiment by about 0.01 eV. Comparing canonical FCI in the TZ2P basis with SHCI in the larger aug-cc-pVQZ basis shows differences of up to 0.158 eV demonstrating that large basis sets are necessary to fully describe correlation in methylene. This was first demonstrated using diffusion Monte Carlo (DMC) results, which are much less sensitive to basis sets and agree with SHCI to within about 0.02 eV. CR-EOMCC(2,3)D relative energies are generally within 1.6 mHa (0.044 eV) of the benchmark SHCI values, indicating that high-level multi-reference coupled cluster calculations are able to correlate six electrons sufficiently to obtain FCI-quality energy gaps.

#### B. Ethylene

Ethylene is another prototypical benchmark system for electronic excitations, including an especially challenging $1^1B_u$ state. Although the $1^1B_u$ state is qualitatively well described by a $\pi-\pi^*$ excitation, a quantitative description requires a thorough accounting of dynamic correlation between $\sigma$ and $\pi$ electrons. Here, SHCI is applied to the low-lying valence states of ethylene: $1^1A_g$, $1^1B_1u$, and $1^3B_1u$, in the ANO-L-pVTZ basis.

Fully correlating ethylene’s twelve valence electrons is a considerably more difficult task than correlating methylene’s six. This is reflected in the fact that SHCI perturbative corrections start to fall below 1.6 mHa only at $\epsilon_V = 7 \cdot 10^{-6}$ Ha (Figure 2). These results suggest that polyatomics with up to twelve valence electrons and triple-zeta basis sets are amenable to treatment at the FCI level using just the variational component of SHCI. Table II compares SHCI total and relative energies with previous FCIQMC and iFCI results. SHCI total energies are only about 1 mHa lower than FCIQMC and the $1^1B_1u - 1^1A_g$ excitation energy is in even better agreement. The $1^3B_1u - 1^1A_g$ energy obtained from iFCI is also in reasonably good agreement, thought it is obtained using a different triple-zeta basis. On the other hand, Table II also indicates that coupled cluster methods must include more than triples excitations in order to obtain SHCI-quality relative energies, as CR-EOMCC(2,3)D results show errors considerably greater than 1.6 mHa (0.044 eV) with respect to the SHCI benchmark values.

The SHCI relative energies support the notion that the vertical excitations cannot be quantitatively compared to the experimental band maxima in ethylene.

Ethylene is the largest system tested for which the perturbative correction is less than 1.6 mHa. This requires using $\epsilon_V = 7 \cdot 10^{-6}$ Ha and 10$^8$ determinants in the variational space, which is near the limit of what can be reasonably stored on contemporary hardware.

As mentioned in Methods: Converting SHCI Energies to the FCI Limit, in larger systems we fit $E_{\text{tot}}$ or $E_{\text{var}}$ to $E_{\text{Fcl}} - E_{\text{var}}$, using a quadratic function and use the fitted function to extrapolate to the no perturbative correction limit, thereby obtaining accurate energies even when the variational energies are not converged. The fits of $E_{\text{tot}}$ and $E_{\text{var}}$ are shown in Fig. 2. The $E_{\text{tot}}$ is nearly flat. To estimate the error of the extrapolation, we performed an additional fit omitting the black points in Fig. 2. The
TABLE I. Methylene/aug-cc-pVQZ total (Ha) and relative (eV) energies

| State | SHCI<sup>a</sup> aug-cc-pVQZ | CR-EOMCC (2,3)<sup>b</sup> aug-cc-pVQZ | FCI<sup>b</sup> TZ2P |
|-------|-------------------------------|------------------------------------|------------------|
| 1^3B<sub>1</sub> | -39.0849(1) | -39.08817 | -39.06674 |
| 1^1A<sub>1</sub> | -39.07404(1) | -39.07303 | -39.04898 |
| 1^1B<sub>1</sub> | -39.03711(1) | -39.03450 | -39.01006 |
| 2^3A<sub>1</sub> | -38.99603(1) | -38.99457 | -38.96847 |

| Gap | SHCI<sup>a</sup> aug-cc-pVQZ | CR-EOMCC (2,3)<sup>b</sup> aug-cc-pVQZ | FCI<sup>b</sup> TZ2P | DMC<sup>d</sup> | Exp<sup>f</sup> |
|-----|-------------------------------|------------------------------------|------------------|---------------|---------|
| 1^3A<sub>1</sub> - 1^3B<sub>1</sub> | 0.393 | 0.412 | 0.483 | 0.406 | 0.400<sup>e</sup> |
| 1^3B<sub>1</sub> - 1^3B<sub>1</sub> | 1.398 | 1.460 | 1.542 | 1.416 | 1.411<sup>e</sup> |
| 2^3A<sub>1</sub> - 1^3B<sub>1</sub> | 2.516 | 2.547 | 2.674 | 2.524 | – |

<sup>a</sup> Using $\epsilon_V = 10^{-5}$ Ha  
<sup>b</sup> FCI/TZ2P results from reference 56  
<sup>c</sup> This work  
<sup>d</sup> Diffusion Monte Carlo results from reference 78  
<sup>e</sup> References 56, 82  
<sup>f</sup> References 56, 83

TABLE II. Ethylene/ANO-L-pVTZ total (Ha) and relative (eV) energies

| State | SHCI<sup>a</sup> ANO-L-pVTZ | CR-EOMCC(2,3)<sup>b</sup> ANO-L-pVTZ | FCIQMC<sup>c</sup> ANO-L-pVTZ | iFCI<sup>d</sup> cc-pVTZ |
|-------|-------------------------------|------------------------------------|------------------|---------------|
| 1^1A<sub>g</sub> | -78.4381(1) | -78.43698 | -78.4370(2) | - |
| 1^1B<sub>1u</sub> | -78.1424(1) | -78.13375 | -78.1407(3) | - |
| 1^1B<sub>1u</sub> | -78.2693(1) | -78.26205 | - | - |

| Gap | SHCI<sup>a</sup> ANO-L-pVTZ | CR-EOMCC(2,3)<sup>b</sup> ANO-L-pVTZ | FCIQMC<sup>c</sup> ANO-L-pVTZ | iFCI<sup>d</sup> cc-pVTZ |
|-----|-------------------------------|------------------------------------|------------------|---------------|
| 1^1B<sub>1u</sub> - 1^1A<sub>g</sub> | 8.05 | 8.25 | 8.06 | - | 7.66<sup>e</sup> |
| 1^3B<sub>1u</sub> - 1^1A<sub>g</sub> | 4.59 | 4.76 | - | 4.64 | 4.3-4.6<sup>f</sup> |

<sup>a</sup> $E_{tot}$ with $\epsilon_V = 7 \cdot 10^{-6}$ Ha  
<sup>b</sup> This work  
<sup>c</sup> FCIQMC/ANO-L-pVTZ results from reference 71  
<sup>d</sup> iFCI/cc-pVTZ results from reference 18  
<sup>e</sup> Experimental band maximum from reference 87  
<sup>f</sup> Experimental band maxima from references 88–90

TABLE III. Comparison of extrapolated ethylene/ANO-L-pVTZ energies obtained using all the points plotted in Fig. 2 with those obtained from omitting the black points.

| State   | $E_{\text{extrap}}$ (Ha) | $E_{\text{extrap}}$ (Ha) |
|---------|--------------------------|--------------------------|
|         | all points               | omit black points        |
| 1^1A<sub>g</sub> | -78.4382                | -78.4385                |
| 1^1B<sub>1u</sub> | -78.1424                | -78.1430                |
| 1^3B<sub>1u</sub> | -78.2693                | -78.2697                |

extrapolated values obtained from the two fits are shown in Table III.

C. Ozone

Ozone’s potential energy surfaces have held great interest due to its role in atmospheric chemistry<sup>93</sup>. An interesting feature predicted by computational studies is the existence of a metastable ring geometry on the ground state surface<sup>57</sup>. A lack of experimental evidence for such a species has fueled multiple studies of the pathway leading to the ring species over the years<sup>92–97</sup>. The most recent such study by Ruedenberg et al. utilizes multi-reference CI with up to quadruple excitations<sup>58</sup>, expending considerable effort on selecting and justifying an active space. To provide an accurate picture at critical points along the theorized pathway with even treatment of all valence electrons, SHCI is applied to ozone’s $2^1A_1$-$1^3A_1$ gap with the cc-pVTZ basis at the three geometries of interest shown in Figure 3: the equilibrium geometry (termed the open
FIG. 2. Fit of variational and total energies of ethylene using all data points. The black line is placed at 1.6 mHa. States are separated from one another by 0.01 Ha for clarity. The tightest SHCI calculation used $\epsilon_V = 7 \times 10^{-4}$ Ha. The extrapolated energies obtained using all data points are compared to those obtained by omitting the black points in Table IV.

FIG. 3. Ozone potential energy surface

TABLE IV. Evolution of ozone $2^1A_1 - 1^1A_1$ gaps (Ha). The RM energy at $\epsilon_V = 2 \times 10^{-4}$ is omitted from the fit.

| $\epsilon_V$ | OM     | TS     | RM    |
|------------|--------|--------|-------|
| $2 \times 10^{-4}$ | 0.1520 | 0.0009 | 0.2897 |
| $1 \times 10^{-4}$ | 0.1522 | 0.0008 | 0.2314 |
| $5 \times 10^{-5}$ | 0.1521 | 0.0005 | 0.2298 |
| $4 \times 10^{-5}$ | 0.1521 | 0.0006 | 0.2293 |
| Extrapolated | 0.1519 | 0.0003 | 0.2254 |

TABLE V. Ozone $2^1A_1 - 1^1A_1$ gaps (eV)

| Geometry | Extrapolated SHCI | MRCl (SDTQ)$^a$ |
|----------|-------------------|-----------------|
| OM       | 4.13              | 3.54-4.63       |
| TS       | 0.01              | 0.05-0.16       |
| RM       | 6.13              | 7.35-8.44       |

$^a$ Reference 58

D. Shorter Polyenes: Butadiene and Hexatriene

Butadiene and hexatriene are part of the polyene series, long studied for their role as prototypical organic conducting polymers. In particular, the spacing of the low-lying valence excited states has proven especially challenging for electronic structure methods. Butadiene and hexatriene are of special interest because their $1^1B_{1u}$ and $2^1A_g$ states are nearly degenerate, resulting in conflicting reports of state ordering at lower levels of theory. In the ANO-L-pVDZ basis, butadiene and hexatriene’s FCI spaces of $10^{26}$ and $10^{38}$ determinants, respectively, are too large for the routine application of FCI-level methods, although limited FCIQMC studies and DMRG studies as well as SHCI ground state calculations have been performed on butadiene. Herein, SHCI is applied to the $1^1A_g$, $1^1B_{1u}$, $1^3B_{1u}$, and $2^1A_g$ states to provide accurate benchmarks and state orderings.
1. Butadiene

Similar to ozone, extrapolation is used to obtain FCI energy estimates for butadiene in the ANO-L-pVDZ basis, as the strongest SHCI calculations at $\epsilon_V = 3 \cdot 10^{-5}$ Ha had perturbative corrections ranging from 12-29 mHa (Figure 4). Besides using orbitals that minimize the SHCI variational energy, for molecules with more than a few atoms a further improvement in the energy convergence can be obtained by localizing the orbitals. Butadiene has $C_{2h}$ symmetry, but the localized orbitals transform as the irreducible representations of the $C_{2h}$ subgroup of $C_{2v}$. Both the $A_g$ and the $B_u$ irreducible representations of $C_{2h}$ transform as the $A'$ representation of $C_{2v}$. Hence calculating the three singlet states, $1^1A_g$, $2^1A_g$ and $1^1B_{2u}$ would require calculating three states simultaneously if localized orbitals are used, and further we would not know if the $2^1A_g$ or the $1^1B_{2u}$ is lower in energy. Consequently, we calculated only the $1^1A_g$ and $1^1B_{2u}$ orbitals using localized orbitals and computed the $2^1A_g$ and $1^1B_{2u}$ states with extended orbitals.

Table VI shows that using the same geometry as in prior FCIQMC,\textsuperscript{23} DMRG,\textsuperscript{12} and iFCI,\textsuperscript{19} calculations leads to a SHCI $1^1A_g$ energy that is 0.4 and 0.9 mHa below the extrapolated DMRG\textsuperscript{12} and iFCI energies respectively. Although FCIQMC has yielded very accurate energies for many systems, in the case of butadiene all three methods (SHCI, iFCI, and DMRG) are in agreement that the FCIQMC energy for the $1^1A_g$ is 8-9 mHa too high. This may be either because of FCIQMC initiation bias or because of an underestimate of the FCIQMC statistical error because of the very long auto-correlation times encountered for large systems. A similar conclusion can be reached for the FCIQMC $1^1B_{2u}$ calculation, as the SHCI energy falls below it by a large amount, 12 mHa.

Turning to relative energies, we see that SHCI is in close agreement with all prior FCI-level theoretical calculations. Both the $1^1B_u-1^1A_g$ and $1^1B_u-1^1A_g$ gaps are 0.08 eV away from the iFCI and FCIQMC values respectively. Although the $2^1A_g-1^1A_g$ gap does not currently have FCI-level benchmarks, the agreement of SHCI’s other relative energies with existing benchmarks supports the accuracy of the extrapolated SHCI value for this gap, which is 6.58 eV. This places the $2^1A_g$ state above $1^1B_u$ in butadiene by 0.13 eV. This small gap is consistent with recent theoretical\textsuperscript{99} and experimental\textsuperscript{101} investigations demonstrating ultrafast population transfer from $1^1B_u$ to $2^1A_g$, which implies close proximity of the two states. As with ethylene, relative energies only qualitatively agree with experiment, supporting prior indications that experimental band maxima of butadiene do not correspond to the vertical excitation energy\textsuperscript{101}.

2. Hexatriene

Hexatriene is at the current frontier of FCI-level computations, with a demanding FCI space of $10^{38}$ determinants in the ANO-L-pVDZ basis. Only one other algorithm, iFCI,\textsuperscript{23} has approached FCI energies for such a large polyatomic - and then only for its singlet-triplet gap. Here we compute the energies of the lowest three singlet states and the lowest triplet state, using values of $\epsilon_V$ as small as $2 \cdot 10^{-5}$ Ha (Figure 5), which results in as many as $9 \cdot 10^7$ variational determinants.

The extrapolated hexatriene energies are reported in Table VII. With the same geometry, SHCI produces a $1^1A_g$ total energy 4 mHa below iFCI. This difference is within the extrapolation uncertainty of SHCI for this
FIG. 5. Extrapolation of hexatriene SHCI energies. States are separated from each other by 0.01 Ha for clarity. The closest SHCI calculation used $\epsilon_V = 2 \times 10^{-5}$ Ha.

TABLE VII. Hexatriene total (Ha) and relative energies (eV)

| State         | Extrapolated SHCI | iFCI     |
|---------------|-------------------|----------|
| $1^1A_g$      | -232.7567(1)      | -232.7527$^c$ |
| $1^3B_u$      | -232.6548(1)      |          |
| $1^1B_u$      | -232.5511(1)      | -         |
| $2^1A_g$      | -232.5517(1)      | -         |

| Gap           | Extrapolated SHCI | CC | iFCI | Exp |
|---------------|-------------------|----|-----|-----|
| $2^1A_g-1^1A_g$ | 5.58              | 5.72$^a$ | -   | 5.21$^o$ |
| $1^1B_u-1^1A_g$ | 5.59              | 5.30$^b$ | -   | 4.95$^e$, 5.13$^e$ |
| $1^3B_u-1^1A_g$ | 2.77              | 2.80$^d$ | 2.81$^d$ | 2.61$^f$ |

- $^a$ CR-EOMCC(2,3)/D/TZVP from reference 67
- $^b$ CCSD(T)/6-31G* from reference 18
- $^c$ iFCI/ANO-L-pVDZ result from reference 19
- $^d$ iFCI/6-31G* result from reference 18
- $^e$ Raman scattering results from reference 110
- $^f$ Electron impact band maximum from reference 111

system. Prior investigations of hexatriene photo dynamics place $1^1B_u$ close in energy to $2^1A_g$. At the vertical excitation geometry, SHCI places $2^1A_g$ below $1^1B_u$ with a small gap of only 0.01 eV. The triplet-singlet $1^3B_u-1^1A_g$ gaps computed by SHCI and iFCI (using the slightly smaller 6-31G* basis) differ by only 0.04 eV. As in the case of butadiene, the SHCI $1^3B_u-1^1A_g$ gap differs significantly from experiment indicating that experimental band maxima do not correspond to vertical excitation energies in hexatriene.

IV. CONCLUSION

SHCI represents an important step forward for SCI methods, providing FCI-quality energies in the largest molecular systems to date. SHCI easily correlates systems of 12 electrons in a triple-zeta basis and can reach FCI-level energies in larger systems through extrapolation. In this paper, CI spaces of $10^7 - 10^8$ determinants were used to effectively handle FCI spaces of $10^{26}$ and $10^{38}$ determinants.

This work has provided new benchmarks and insights for the valence states of some commonly investigated molecular systems. Specifically, high-quality SHCI energetics for ozone give strong evidence that the theoretical RM structure has a significant barrier to relaxation, and thus should be observable by experiment. Investigation of butadiene and hexatriene lead to the highest level $2^1A_g-1^1B_u$ state-orderings in these systems to date, placing $2^1A_g$ above $1^1B_u$ in butadiene, and minutely below $1^1B_u$ in hexatriene. In short, SHCI has shown itself to be an efficient means of obtaining FCI-level energetics, and we look forward to it providing physical insight into other chemically interesting systems with up to dozens of electrons.

ACKNOWLEDGMENTS

ADC and PMZ would like to thank David Braun for computational support and the University of Michigan for financial support. SS acknowledges the startup package from the University of Colorado. AAH, MO and CJU were supported in part by NSF grant ACI-1534965. Some of the computations were performed on the Bridges computer at the Pittsburgh Supercomputing Center, supported by NSF award number ACI-1445606.

1. P. Knowles and N. Handy, Chem. Phys. Lett. 111, 315 (1984).
2. J. Olsen, B. O. Roos, P. Jørgensen and H. J. A. Jensen, J. Chem. Phys. 89, 2185 (1988).
3. J. Olsen, P. Jørgensen and J. Simons, Chem. Phys. Lett. 169, 463 (1990).
4. J. Olsen, O. Christiansen, H. Koch and P. Jørgensen, J. Chem. Phys. 105, 5082 (1996).
5. E. Rossi, G. L. Bendazzoli, S. Evangelisti and D. Maynau, Chem. Phys. Lett. 310, 530 (1999).
6. A. Dutta and C. D. Sherrill, J. Chem. Phys. 118, 1610 (2003).
7. Z. Gan, D. J. Grant, R. J. Harrison and D. A. Dixon, J. Chem. Phys. 125, 124311 (2006).
8. S. R. White, Phys. Rev. B 48, 10345 (1993).
9. S. R. White and R. L. Martin, J. Chem. Phys. 110, 4127 (1999).
10. G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. 116, 4462 (2002).
11. G. K.-L. Chan and S. Sharma, Annu. Rev. Phys. Chem.
An extrapolated value of -155.5578 was obtained by fitting the values in Table VI of Ref. [12] to either a linear or a quadratic function of $1/M$, where $M$ is the DMRG bond dimension.