Potential energy surface for interactions between two hydrogen molecules

Konrad Patkowski,[1,a)] Wojciech Cencek,[1] Piotr Jankowski,[2] Krzysztof Szalewicz,[1] James B. Mehl,[3] Giovanni Garberoglio,[4] and Allan H. Harvey[5]
[1]*Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, USA*
[2]*Department of Quantum Chemistry, Institute of Chemistry, Nicolaus Copernicus University, Gagarina 7, PL-87-100 Toruń, Poland*
[3]*P.O. Box 307, Orcas, Washington 98280, USA*
[4]*CNISM and Dipartimento di Fisica, Università di Trento, Via Sommarive 14, 38100 Trento, Italy*
[5]*Physical and Chemical Properties Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA*

(Received 8 July 2008; accepted 4 August 2008; published online 4 September 2008)

Nonrelativistic clamped-nuclei energies of interaction between two ground-state hydrogen molecules with intramolecular distances fixed at their average value in the lowest rovibrational state have been computed. The calculations applied the supermolecular coupled-cluster method with single, double, and noniterative triple excitations [CCSD(T)] and very large orbital basis sets—up to augmented quintuple zeta size supplemented with bond functions. The same basis sets were used in symmetry-adapted perturbation theory calculations performed mainly for larger separations to provide an independent check of the supermolecular approach. The contributions beyond CCSD(T) were computed using the full configuration interaction method and basis sets up to augmented triple zeta plus midbond size. All the calculations were followed by extrapolations to complete basis set limits. For two representative points, calculations were also performed using basis sets with the cardinal number increased by one or two. For the same two points, we have also solved the Schrödinger equation directly using four-electron explicitly correlated Gaussian (ECG) functions. These additional calculations allowed us to estimate the uncertainty in the interaction energies used to fit the potential to be about 0.15 K or 0.3% at the minimum of the potential well. This accuracy is about an order of magnitude better than that achieved by earlier potentials for this system. For a near-minimum T-shaped configuration with the center-of-mass distance $R=6.4$ bohrs, the ECG calculations give the interaction energy of -56.91 ± 0.06 K, whereas the orbital calculations in the basis set used for all the points give -56.96 ± 0.16 K. The computed points were fitted by an analytic four-dimensional potential function. The uncertainties in the fit relative to the *ab initio* energies are almost always smaller than the estimated uncertainty in the latter energies. The global minimum of the fit is −57.12 K for the T-shaped configuration at $R=6.34$ bohrs. The fit was applied to compute the second virial coefficient using a path-integral Monte Carlo approach. The achieved agreement with experiment is substantially better than in any previous work. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2975220]

I. INTRODUCTION

Thermophysical properties of hydrogen gas are important in many research areas as well as for industrial applications. Measured values of quantities such as virial coefficients, viscosities, and thermal conductivities are available. The experimental techniques are well developed and the uncertainties in measured quantities have been brought to near the minimal values possible with the current technology. Further significant reduction in these uncertainties is not likely. The same quantities can also be obtained from statistical mechanics calculations based on *ab initio* intermolecular potentials. For helium, this approach led to theoretically predicted thermophysical properties that are significantly more accurate than the measured ones.[1–15] Since the hydrogen molecule has only two electrons, similar to the helium atom, the *ab initio* approach may be a fruitful route in this case as well. However, *ab initio* calculations of the $(H_2)_2$ interaction potential are more difficult than those for He_2. Even for a single configuration, the computational effort at a given level of theory and basis set is larger for $(H_2)_2$ due to the twice larger size of the basis set and lower symmetry of the system. The calculation of the potential is even harder since the He–He potential is only one-dimensional whereas that for H_2-H_2 is six-dimensional (four-dimensional if rigid monomers are assumed as in the present work). Thus, it is unlikely that one can achieve the He_2-level accuracy of Ref. 12 in the case of the H_2-H_2 potential. However, matching or exceeding the current accuracies of the measured thermodynamic properties may be within reach.

Knowledge of an accurate H_2-H_2 potential is also very important for simulations of *para*-hydrogen clusters in he-

[a)]Electronic mail: patkowsk@udel.edu.

lium nanodroplets. These simulations are motivated by experimental indications that molecular hydrogen may exhibit superfluidity in nanodroplet environments[16] or in small clusters.[17] These findings are supported by theoretical studies;[17–19] see Ref. 20 for a recent review.

There exists a large number of published H_2-H_2 interaction potentials. Until the very recent work of Hinde[21] (see Sec. VI), the four-dimensional potential of Diep and Johnson[22] was believed to be the most accurate. The calculations of Diep and Johnson[22] used the coupled-cluster method with single, double, and noniterative triple excitations [CCSD(T)] and orbital basis sets up to augmented quadruple zeta quality. The results were extrapolated to the complete basis set (CBS) limit using the values calculated in the aug-cc-pVTZ and aug-cc-pVQZ bases.[23,24] No estimate of accuracy was given by Diep and Johnson.[22] The changes in the potential due to the corrections introduced in the Erratum of Ref. 22 are of the order of 2 K or 4% near the minimum. The room-temperature virial coefficient computed from the original Diep–Johnson (DJ) potential[25] deviates from the measured value by more than the experimental uncertainties, and the agreement is worse at lower temperatures. Thus, the accuracy of this potential is clearly far from that needed for replacement of measured thermophysical properties by theoretically computed ones. As we will discuss later on, one reason for this problem could be some apparent inaccuracies in the CCSD(T) interaction energies computed by Diep and Johnson.[22]

The system of four hydrogen atoms is also a frequently used model for investigating the dynamics of chemical reactions. There have been a large number of H_4 potential energy surfaces developed for this purpose. Such surfaces always include the intramonomer coordinates and often are computed also for excited electronic states of the system. However, these surfaces are of low accuracy in the van der Waals region of interest to the present work since low-level methods and small basis sets have been used so far. For example, the most elaborate such work to date[26] used a multireference configuration interaction (CI) method with double excitations only and a $[4s3p1d]$ basis set. The surface of Ref. 26 was used in close-coupling calculations of rotationally inelastic cross sections in H_2-H_2 collisions and gave rate coefficients an order of magnitude too small.[27] In contrast, the DJ surface gave a reasonable agreement with experiment.

Once an accurate H_2-H_2 potential is known, one can in principle calculate many important thermophysical properties of hydrogen, such as the second virial coefficient and the low-density limits of the viscosity and thermal conductivity. However, such calculations are also significantly more difficult than in the case of helium. Similar to helium, hydrogen is a highly quantum system. Thus, classical or semiclassical calculations of these properties would be inadequate except at high temperatures (room temperature and above). For helium, it is possible to use exact quantum formulas that require knowledge of (quantum) scattering cross sections in a broad range of scattering energies. Calculations of such extensive scattering data for H_2-H_2 would be a daunting task. For an example of state-of-the-art close-coupling calculations for H_2-H_2, see Ref. 27. For heavier linear molecules, quantum scattering calculations needed for thermophysical properties can be adequately replaced by classical-trajectory calculations (see, for example, Refs. 28 and 29), and for H_2-H_2 this approach can be used at high temperatures. Some properties, such as the second virial coefficient, can be extracted from path-integral Monte Carlo (PIMC) calculations for two hydrogen molecules. However, it is not clear if this method can be applied to other properties of interest such as viscosity or thermal conductivity. For the virial coefficient, one can also use the effective potential approximation[30,31] which should give an accurate account of the quantum effects at moderate and high temperatures.

We present here a development of a new four-dimensional potential for H_2-H_2. The hydrogen molecules were assumed to be rigid at their average bond length in the ground rovibrational state. The assumption of rigidity was needed since it would not be feasible at present to perform calculations of interaction energies with the desired level of accuracy at the number of configurations necessary to develop a six-dimensional potential (the number of configurations would have to be at least an order of magnitude larger than that in the case of a four-dimensional potential). Moreover, calculations of the scattering cross sections with a six-dimensional potential would be very difficult. One may hope that the inclusion of intramonomer vibrational motion is not very important for hydrogen as the energy of the lowest vibrational excitation is high (5987 K, Ref. 32) compared to thermal energies of interest. For a model study of the vibrational couplings in H_2-H_2, see Ref. 33. A more important monomer-flexibility effect could be the stretch of the H–H bond in rotational excitations of H_2 molecules. The latter effect could be investigated by computing a four-dimensional potential with one of the hydrogen molecules frozen at the average separation in the ground vibrational and the first excited rotational state. One can also compute a six-dimensional surface at a lower accuracy level than that in the present calculations and use it to determine corrections to thermophysical properties due to monomer flexibility.

The initial aim of this work was to obtain a potential with an uncertainty below 0.5 K or 1% in the region of the potential minimum and to reach the same percentage accuracy for larger intermolecular separations. For smaller separations, the absolute uncertainties on the negative part of the repulsive wall and for small positive potential values should preferably be similar to those near the minimum. The uncertainties in the strongly repulsive region should be below 1%. Realization of these goals would increase the accuracy of the H_2-H_2 potential by about an order of magnitude compared to the DJ potential.[22] We will demonstrate that our initial accuracy goals were actually surpassed by our calculations.

II. ORBITAL CALCULATIONS

The geometry of a dimer composed of two rigid hydrogen molecules (with the intramonomer distances frozen at the average value in the ground rovibrational state, i.e., at 1.448 736 bohrs,[34,35] where 1 bohr $\approx 0.529\,177\,2 \times 10^{-10}$ m) is defined by four coordinates: the center-of-mass (COM) separation R and three angles θ_1, θ_2, and ϕ. In

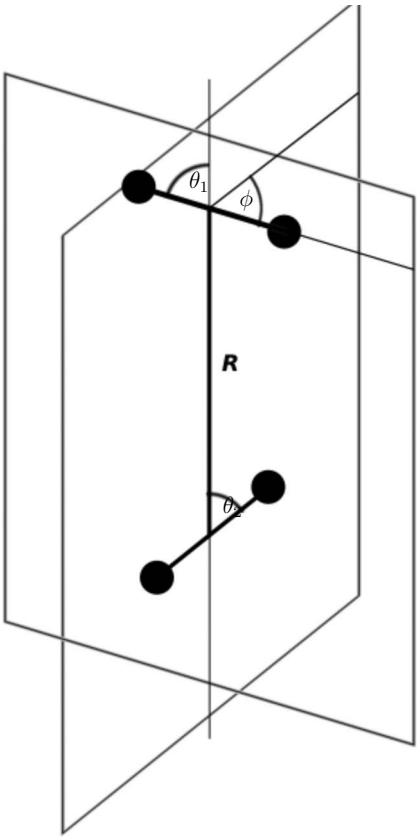

FIG. 1. Geometry of the H_2 dimer.

the coordinate system where both monomers' COMs lie on the z axis, θ_1 and θ_2 are defined as the angles between the positive half of the z axis and the H–H bond line of the first and second monomer, respectively, whereas ϕ is the dihedral angle between the planes containing the z axis and one of the monomers, see Fig. 1. Because of symmetry, it is sufficient to consider the values (all angles are given in degrees here and throughout the rest of the work) $0 \leq \theta_2 \leq \theta_1 \leq 90$ and $0 \leq \phi \leq 180$. Moreover, whenever $\theta_2 = 0$, the angle ϕ becomes arbitrary and it will be set to $\phi = 0$. Finally, if $\theta_1 = 90$, the configurations obtained for ϕ and $180 - \phi$ are equivalent for any θ_2, and it is enough to assume that $0 \leq \phi \leq 90$ in this case.

For a given R, the basic angular grid used in our calculations consists of all $(\theta_1, \theta_2, \phi)$ triplets such that $\theta_1, \theta_2 \in \{0, 30, 60, 90\}$, $\phi \in \{0, 45, 90, 135, 180\}$, and the symmetry-related conditions defined in the previous paragraph are satisfied. This grid contains 28 symmetry-distinct angular configurations. Our main batch of grid points was formed by $5 \times 28 = 140$ configurations corresponding to $R = 4.0$, 5.0, 6.0, 7.0, and 8.0 bohrs. The grid used in our final fit consists of 213 points: 140 points from the main batch, the near-minimum geometry $(R, \theta_1, \theta_2, \phi) = (6.4, 90, 0, 0)$, 67 additional points along the radial cuts $(\theta_1, \theta_2, \phi) = (0, 0, 0)$, (90,0,0), (90,90,0), and (90,90,90) placed every 0.25 bohr such that $3.0 \leq R \leq 9.0$ bohrs and the calculated energy does not exceed 5000 K, and 5 points along the angular cut $(R, \theta_1, \theta_2) = (6.5, 90, 90)$, $\phi = 15, 30, 45, 60, 75$. For the purpose of testing the quality of our analytical fit, we have also calculated some *ab initio* points in the asymptotic regime, as well as 20 geometries $\boldsymbol{G}_1, \ldots, \boldsymbol{G}_{10}, \boldsymbol{G}'_1, \ldots, \boldsymbol{G}'_{10}$ chosen such that $\boldsymbol{G}_n$ $(\boldsymbol{G}'_n)$ is as far away as possible from all grid points (from all main-batch points) and from all $\boldsymbol{G}_i$ $(\boldsymbol{G}'_i)$, $i = 1, \ldots, n-1$, with an additional restriction $4.0 \leq R \leq 9.0$ bohrs. For this purpose, the distance $d(\boldsymbol{C}_i, \boldsymbol{C}_j)$ between two configurations (quadruplets of Cartesian coordinates of atoms) $\boldsymbol{C}_i = \{\boldsymbol{P}_{i1}, \boldsymbol{P}_{i2}, \boldsymbol{P}_{i3}, \boldsymbol{P}_{i4}\}$ and $\boldsymbol{C}_j = \{\boldsymbol{P}_{j1}, \boldsymbol{P}_{j2}, \boldsymbol{P}_{j3}, \boldsymbol{P}_{j4}\}$, where the atom pairs at $(\mathbf{P}_{i1}, \mathbf{P}_{i2})$, $(\mathbf{P}_{i3}, \mathbf{P}_{i4})$, $(\mathbf{P}_{j1}, \mathbf{P}_{j2})$, and $(\boldsymbol{P}_{j3}, \boldsymbol{P}_{j4})$ are the ones that form covalent bonds, is defined as

$$d(\boldsymbol{C}_i, \boldsymbol{C}_j) = \sqrt{\sum_{k=1}^{4} (d_{ik} - d_{jk})^2}, \tag{1}$$

where the quadruplets $(d_{i1}, d_{i2}, d_{i3}, d_{i4})$ and $(d_{j1}, d_{j2}, d_{j3}, d_{j4})$ are composed of the nonbonded interatomic distances $\{|\boldsymbol{P}_{i1}\boldsymbol{P}_{i3}|, |\boldsymbol{P}_{i1}\boldsymbol{P}_{i4}|, |\boldsymbol{P}_{i2}\boldsymbol{P}_{i3}|, |\boldsymbol{P}_{i2}\boldsymbol{P}_{i4}|\}$ and $\{|\boldsymbol{P}_{j1}\boldsymbol{P}_{j3}|, |\boldsymbol{P}_{j1}\boldsymbol{P}_{j4}|, |\boldsymbol{P}_{j2}\boldsymbol{P}_{j3}|, |\boldsymbol{P}_{j2}\boldsymbol{P}_{j4}|\}$, respectively, sorted in ascending order. All CCSD(T) calculations employed the MOLPRO code.[36,37] The full configuration interaction (FCI) calculations were done using the LUCIA program[38] and the symmetry-adapted perturbation theory (SAPT) calculations using the SAPT2006 codes.[39]

The computational resources needed to perform calculations for a given geometry, especially at the FCI level of theory, vary greatly with the point-group symmetry of the dimer. For different angular configurations, the Abelian spatial symmetry of the dimer varies from C_1 (no symmetry) to D_{2h} (eightfold symmetry). Note that in the latter case the symmetry of a system composed of one monomer and ghost centers in place of the other monomer's atoms, needed for counterpoise[40] calculations, is lower than D_{2h}; however, this fact is of no practical importance since the dimer calculation is the time- and memory-limiting step. The calculations performed for all the grid points had to employ basis sets small enough to be feasible without any point-group symmetry present. Consequently, the largest bases determining the potential were aug-cc-pV5Z+(33221) at the CCSD(T) level and aug-cc-pVTZ+(332) at the FCI level. The numbers in parentheses denote basis functions placed on the intermolecular bond (halfway between the monomer COMs). The $(33221) \equiv (3s3p2d2f1g)$ set is composed of primitive Gaussian functions with exponents equal to 0.9, 0.3, and 0.1 for sp, 0.6 and 0.2 for df, and 0.35 for g symmetry, and the (332) set is the $(3s3p2d)$ subset of (33221). On the other hand, for symmetric configurations we were able to perform calculations with significantly larger basis sets. A comparison of the CCSD(T) and FCI interaction energies calculated in various bases is presented in Table I for the near-minimum T-shaped configuration (C_{2v} symmetry) with $R = 6.4$ bohrs and in Table II for the linear configuration with $R = 6.0$ bohrs. Apart from the aug-cc-pVXZ (denoted aXZ for short) basis sets of Dunning *et al.*, with or without midbond functions, we have employed doubly augmented bases, dXZ=d-aug-cc-pVXZ (however, the dXZ results are incomplete because of the linear dependencies encountered), as well as the modified Dunning-type bases am$X$Z $\equiv$ aug-mcc-pVXZ constructed (for hydrogen only) in Refs. 41 and 42. The bases am$X$Z, of the same sizes as regular aXZ, were intended to form a more regular sequence than the original bases (in order to facilitate

TABLE I. Comparison of the CCSD(T) and FCI interaction energies (in K) for the minimum geometry, which is a T-shaped configuration $(\theta_1,\theta_2,\phi)=(90,0,0)$ with $R=6.4$ bohrs and the intramonomer distance of 1.448 736 bohrs, calculated and CBS-extrapolated with the X^{-3} formula from various basis sets. The values obtained using the basis sets employed for the entire surface have been marked in boldface.

Basis	Size	CCSD(T)	CCSD(T)/CBS	FCI	δ_{FCI}	δ_{FCI}/CBS
aDZ	36	−34.99		−36.87	−1.89	
aTZ	92	−50.46	−55.86	−52.44	−1.97	−2.01
aQZ	184	−53.70	−55.84	−55.61	−1.91	−1.86
a5Z	320	−54.41	−55.21			
a6Z	508	−54.76	−55.21			
aDZ+(332)	58	−53.14		−55.15	−2.01	
aTZ+(332)	114	−53.91	−54.67	−55.87	−1.96	**−1.94**
aQZ+(332)	206	−54.63	−54.99	−56.52	−1.89	−1.84
a5Z+(332)	342	−54.85	−55.04			
a6Z+(332)	530	−54.95	−55.10			
aQZ+(33221)	229	−54.77				
a5Z+(33221)	365	−54.91	**−55.02**			
a6Z+(33221)	553	−54.99	−55.10			
dTZ	128	−51.02		−53.00	−1.98	
dQZ	248	−54.06	−56.06			
d5Z	420	−54.63	−55.29			
d6Z	652	−54.90	−55.24			
dTZ+(332)	150	−54.04		−56.00	−1.96	
dQZ+(332)	270	−54.69	−55.03			
d5Z+(332)	442	−54.89	−55.06			
d6Z+(332)	674	−54.98	−55.11			
dQZ+(33221)	293	−54.80				
d5Z+(33221)	465	−54.94	−55.05			
d6Z+(33221)	697	−55.01	−55.09			
amTZ	92	−50.19				
amQZ	184	−53.63	−55.86			
am5Z	320	−54.41	−55.29			
am6Z	508	−54.75	−55.21			
amTZ+(332)	114	−53.89				
amQZ+(332)	206	−54.62	−54.99			
am5Z+(332)	342	−54.85	−55.06			
am6Z+(332)	530	−54.96	−55.10			
amQZ+(33221)	229	−54.76				
am5Z+(33221)	365	−54.91	−55.05			
am6Z+(33221)	553	−54.99	−55.10			

CBS extrapolations). However, as Tables I and II show, the differences between the am*X*Z and a*X*Z interaction energies of the H_2-H_2 dimer, both computed for a given X and extrapolated, are virtually negligible. The bases am*X*Z are available for X up to 8 (Ref. 43), whereas the regular a*X*Z bases are only available up to $X=6$. For the linear configuration, we were able to utilize the $X=7$ set at the CCSD(T) level, and Table II presents results in bases containing as many as 801 functions, whereas for the T-shaped minimum configuration, we employed basis sets up to $X=6$, the largest of those containing 697 functions. At the FCI level, the largest bases we were able to use contained 229 and 206 functions, respectively.

The CBS extrapolation technique employed is the most popular and has been supported by numerous analytical and numerical studies.[7,44–49] Specifically, all post-self-consistent field (SCF) contributions to the interaction energy, i.e., the correlation contribution to the CCSD(T) interaction energy,

$$\delta E_{\mathrm{int}}^{\mathrm{CCSD(T)}} = E_{\mathrm{int}}^{\mathrm{CCSD(T)}} - E_{\mathrm{int}}^{\mathrm{SCF}}, \tag{2}$$

or the contribution beyond CCSD(T),

$$\delta E_{\mathrm{int}}^{\mathrm{FCI}} = E_{\mathrm{int}}^{\mathrm{FCI}} - E_{\mathrm{int}}^{\mathrm{CCSD(T)}}, \tag{3}$$

where $E_{\mathrm{int}}^{\mathrm{T}}$ denotes the interaction energy computed at a T level of theory, have been extrapolated according to the X^{-3} formula, i.e., we assumed that the energies $E(X)$ calculated in X-tuple zeta basis sets approach their CBS limit $E(\infty)$ like $E(\infty)+AX^{-3}$ for some constant A. Since the SCF interaction energies exhibit much faster, exponential convergence to the CBS limits,[50] these energies were not extrapolated. All the extrapolated energies listed in Tables I and II for a given X-tuple zeta basis set have been obtained as described above from energies calculated in this basis set and in a basis set of the same type with X lower by one (e.g., the result given in the a6Z row has been extrapolated from the values calculated in bases a5Z and a6Z). All the supermolecular interaction energies were computed in the counterpoise-corrected way to

TABLE II. Comparison of the CCSD(T) and FCI interaction energies (in K) for the linear geometry $(\theta_1,\theta_2,\phi)=(0,0,0)$ with $R=6.0$ bohrs and the intramonomer distance of 1.448 736 bohrs calculated and CBS-extrapolated with the X^{-3} formula from various basis sets. The values obtained using the basis sets employed for the entire surface have been marked in boldface. Some doubly augmented bases have not been included because of CC convergence problems.

Basis	Size	CCSD(T)	CCSD(T)/CBS	FCI	δ_{FCI}	δ_{FCI}/CBS
aDZ	36	55.01		51.00	−4.01	
aTZ	92	44.98	37.89	41.03	−3.95	−3.93
aQZ	184	40.10	37.60	36.36	−3.74	−3.59
a5Z	320	39.12	38.29			
a6Z	508	38.73	38.19			
aDZ+(332)	58	37.35		33.30	−4.05	
aTZ+(332)	114	39.49	39.36	35.63	−3.86	**−3.78**
aQZ+(332)	206	39.06	38.68	35.35	−3.71	−3.60
a5Z+(332)	342	38.75	38.51			
a6Z+(332)	530	38.58	38.36			
aQZ+(33221)	229	38.87		35.18	−3.69	
a5Z+(33221)	365	38.65	**38.51**			
a6Z+(33221)	553	38.53	38.36			
dTZ	128	44.14				
dQZ	248	39.73	37.47			
dTZ+(332)	150	39.53				
dQZ+(332)	270	39.00	38.61			
dQZ+(33221)	293	38.85				
amTZ	92	45.18				
amQZ	184	40.17	37.59			
am5Z	320	39.17	38.30			
am6Z	508	38.75	38.17			
am7Z	756	38.55	38.23			
amTZ+(332)	114	39.73				
amQZ+(332)	206	39.06	38.60			
am5Z+(332)	342	38.76	38.49			
am6Z+(332)	530	38.58	38.36			
am7Z+(332)	778	38.49	38.34			
amQZ+(33221)	229	38.87				
am5Z+(33221)	365	38.65	38.50			
am6Z+(33221)	553	38.53	38.37			
am7Z+(33221)	801	38.46	38.36			

remove the basis set superposition error.[4,40,51–53]

The values in Table I indicate that the CBS limits for the CCSD(T) interaction energy and for the $\delta E_{\mathrm{int}}^{\mathrm{FCI}}$ contribution are close to −55.09 and −1.84 K, respectively, for the minimum geometry. For the linear configuration at $R=6$ bohrs (Table II), these contributions amount to 38.36 and −3.60 K, respectively. These values should be compared with the ones obtained using basis sets employed for the entire potential energy surface, set in boldface in Tables I and II and referred to below as "whole-grid" values. Thus, judging from our best CBS-limit estimations, our CCSD(T) component of the potential is in error by about 0.07 K at the minimum and 0.15 K for the linear configuration considered. The corresponding uncertainties in the $\delta E_{\mathrm{int}}^{\mathrm{FCI}}$ term are about 0.10 and 0.18 K, respectively. One can clearly see from the results in Tables I and II that the CBS-extrapolated whole-grid values are much closer to the best-estimate CBS limits (inferred from calculations in larger basis sets than used for the whole grid) than the nonextrapolated results obtained using the same basis sets both for the CCSD(T) and $\delta E_{\mathrm{int}}^{\mathrm{FCI}}$ contributions, even though in the latter case the extrapolation involves results calculated in a very small aDZ+(332) basis set. Thus, we have used the extrapolated CCSD(T) and FCI contributions for our final H_2 dimer potential surface. Note that, for the configurations investigated in Tables I and II, the basis set incompleteness errors at the CCSD(T) and FCI levels exhibit strong mutual cancellation. Even without this cancellation, however, the errors of our potential are well within our 0.5 K goal in the minimum region.

When larger-basis results are not available, a simple estimate of the basis set incompleteness error at a given level of theory can be obtained as the magnitude of the difference between the value extrapolated from $(X-1)$-tuple zeta and X-tuple zeta bases (i.e., our final approximation to the interaction energy at this level of theory) and the value calculated in the X-tuple zeta basis set. As shown in Ref. 49, such an estimate is usually very conservative, and often the true uncertainties are much smaller. The respective error estimates for the $E_{\mathrm{int}}^{\mathrm{CCSD(T)}}$ and $\delta E_{\mathrm{int}}^{\mathrm{FCI}}$ terms in the whole-grid basis (values given in boldface in Tables I and II) obtained in this way amount to 0.11 and 0.02 K, respectively, at the minimum and to 0.14 and 0.08 K, respectively, at the linear geometry.

TABLE III. Comparison of the H_2-H_2 interaction energies computed in the basis aug-cc-pVDZ in the present work with those of Ref. 22. The calculations were performed for the T-shaped configuration with $R=3.4$ Å and the intramonomer distance of 1.449 bohrs (as used in Ref. 22). All energies are in kelvins.

Method	Present	Reference 22
MP2	−28.155	−28.135
MP4	−37.914	−37.924
CCSD(T)	−35.445	−37.987
FCI	−37.292	−38.050

Comparison with the CBS-limit energies shows that whereas the error estimates at the CCSD(T) level are very reasonable, the $\delta E_{\rm int}^{\rm FCI}$ errors are underestimated. This fact may indicate that the true convergence pattern of the $\delta E_{\rm int}^{\rm FCI}$ term might be, at least for some distances, closer to AX^{-1} or AX^{-2} than to AX^{-3}. Such a different convergence behavior has indeed been observed[49] for some components of the helium dimer interaction energy, at least for intermonomer distances around the van der Waals minimum. There is, however, no conclusive evidence supporting any particular choice of the exponent α in $AX^{-\alpha}$. Therefore, we have used the conventional value of $\alpha=3$ to extrapolate $\delta E_{\rm int}^{\rm FCI}$ as well as $\delta E_{\rm int}^{\rm CCSD(T)}$ but, to be on the safe side, we arbitrarily enlarged the uncertainty in the former term by a factor of 2. In other words, we have chosen $\sigma=\sigma_{\rm CCSD(T)}+2\sigma_{\rm FCI}$ (added linearly), where $\sigma_{\rm CCSD(T)}$ and $\sigma_{\rm FCI}$ are the differences between the extrapolated and calculated results, as specified above, as the error estimate for our *ab initio* points. Since our error bars are not rigorous in any event, this level of confidence should be adequate. The error estimates obtained in this way are about 0.15 K or 0.3% at the minimum of the potential. These estimates do not exceed 0.5 K in the negative-energy region, 1 K in the positive-energy region below 100 K, and 0.7% on the repulsive wall above 100 K. In the strongly repulsive region above 1000 K, the relative uncertainties are below 0.3%. For R larger than the distances of the radial minimum, the uncertainties are usually smaller than 0.3% except for near-linear configurations where such uncertainties are smaller than 0.07 K. The larger relative uncertainties in the latter case are due to strong cancellations between the electrostatic and dispersion energies for such configurations. The interaction energies and error estimates for the potential grid points and for all additional points are given in the supporting information.[54]

A numerical comparison with the results of Diep and Johnson[22] is difficult since these authors included in their paper the computed values only for one point (all other results were given only in figures). Table III compares our computed values for the same dimer configuration (T shaped with $R=3.4$ Å, using Diep and Johnson's[22] intramonomer distance of 1.449 bohrs), the same levels of theory, and the same aug-cc-pVDZ basis set as used in Ref. 22. The results obtained using many-body perturbation theory with the Møller–Plesset (MP) decomposition of the Hamiltonian through the second (MP2) and fourth (MP4) orders agree to within 0.02 K. The reason for the small discrepancies could be different integral thresholds in the calculations of Diep and Johnson[22] or perhaps a slightly different geometry (at the same time, these results show that the effects of possible differences in the geometry are minor). However, for the CCSD(T) and FCI energies, the discrepancies are dramatic, amounting to 2.54 and 0.76 K, respectively. The reasons for the discrepancies are unknown. We can only speculate that the convergence thresholds in the calculations of Diep and Johnson[22] might have been significantly too large. We have checked our CCSD(T) result by computing it independently with the MOLPRO (Ref. 36) and DALTON (Ref. 55) packages and the results were identical to all digits given in Table III. Thus, in addition to errors of a couple percent due to the basis set and theory level truncation, Diep and Johnson's[22] results may have some random errors due to the numerical inaccuracy of their CCSD(T) calculations.

More extensive comparisons are possible with the fitted potential of Diep and Johnson.[22] The interaction energies obtained from our *ab initio* calculations, from our analytical fit described in Sec. IV, and from the fit of Diep and Johnson[22] are plotted in Figs. 2–5 as functions of R for the angular configurations $(\theta_1,\theta_2,\phi)=(0,0,0)$, (90,0,0), (90,90,0), and (90,90,90), respectively. Another one-dimensional cut of the potential, along the ϕ coordinate for $R=6.5$ bohrs and $\theta_1=\theta_2=90$, is presented in Fig. 6 (note that considering only $0\leq\phi\leq90$ is sufficient in this case because of symmetry). The uncertainties in our *ab initio* points are always smaller than the sizes of the squares representing these points [except for the point (5.5,90,0,0) in Fig. 3 where the error bars are slightly larger than the size of the square]. Figures 2–6 show that our potential is noticeably deeper than the potential obtained by Diep and Johnson,[22] although, interestingly, the difference between the two potentials is very small for the minimum angular configuration (90,0,0). To explain the differences between our results and those of Diep and Johnson,[22] we have performed calculations at the theory level used by Diep and Johnson[22] [i.e., CCSD(T) in bases aug-cc-pVTZ and aug-cc-pVQZ without bond functions, followed by the CBS extrapolations of the total interaction energies according to the X^{-3} scheme] for all our grid points. These DJ-level energies are also plotted in Figs. 2–6 together with their error bars inferred from the CBS extrapolation in the same way as for our energies. The resulting (pointwise) potential is indeed shallower than our highest-level one, although in most cases the DJ-level grid-point energies are closer to our results than the energies predicted by Diep and Johnson's[22] analytical fit. For some configurations, in particular, for (90,90,0), the uncertainties in this fit are quite substantial, on the order of 1 K. We have chosen $R=6.0$ and 7.0 bohrs, two distances from the general region of radial minima, and computed the differences between the DJ-level *ab initio* results (from values calculated by us) and our highest-level ones for all 28×2 angular configurations. These differences are highly anisotropic and amount to 0.51–1.81 and 0.75–1.10 K for $R=6.0$ and 7.0 bohrs, respectively, depending on the angular configuration. These values are actually smaller than the effects beyond the CCSD(T) level of theory, neglected by Diep and Johnson[22] and estimated by us to range from −1.94 to −3.78 K for $R=6.0$ bohrs and from −0.83 to −1.51 K for $R=7.0$ bohrs. Thus, the DJ-level results exhibit some fortuitous cancellation of errors between

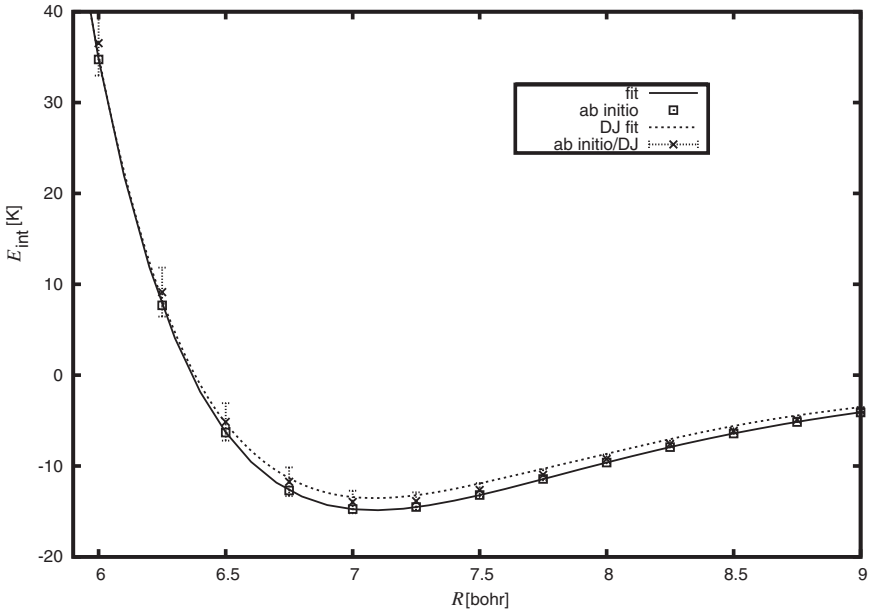

FIG. 2. Calculated and extrapolated (squares) and fitted (solid line) H_2-H_2 interaction energies for the angular configuration $(\theta_1,\theta_2,\phi)=(0,0,0)$ plotted as functions of R. The DJ potential, given in the Erratum of Ref. 22 (labeled "DJ fit"), and the DJ-level *ab initio* results and their uncertainties (obtained in the way described in the text and labeled by "x") are also drawn for comparison.

the post-CCSD(T) contribution and the basis set incompleteness effects at the CCSD(T) level. Despite this cancellation, the DJ-level energies differ from our results by up to 5–6% even for $R=8.0$ bohrs, where the uncertainties in our highest-level potential are more than an order of magnitude smaller than these differences. For some orientations, in particular, for (90,0,0), the inaccuracies in the fit provide additional cancellations, leading to near overlap of the DJ fit with our fit in Fig. 3.

Our experience from the helium dimer calculations[12] shows that, for large intermonomer separations, the accuracy of the CCSD(T)/FCI approach employed here is matched or even surpassed by the accuracy of an approach utilizing SAPT interaction energies[56,57] supplemented by the FCI correction. Therefore, we have checked the results for all grid points with $R=8.0$ bohrs, as well as for additional points along the linear and angular cuts presented in Figs. 2–6, by performing SAPT calculations using the same basis sets as for the CCSD(T) method. The SAPT interaction energies were calculated at the conventional second-order level. The residual Hartree–Fock correction was omitted as recommended for dimers composed of nonpolar monomers,[58] i.e., the quantity $E^{[2]}_{\mathrm{SAPT}}$, defined, e.g., in Ref. 59, was chosen as the SAPT result. The final extrapolated energy was then obtained in the same way as for the CCSD(T)/FCI approach, i.e., as a sum of the SCF-level SAPT energy calculated in basis aug-cc-pV5Z+(33221), the correlated-level SAPT energy extrapolated from bases aug-cc-pVQZ+(33221) and aug-cc-pV5Z+(33221), and the FCI contribution missing in SAPT, extrapolated from bases aug-cc-pVDZ+(332) and aug-cc-pVTZ+(332).

We have estimated the uncertainties in the total energy (including the FCI correction) obtained in the SAPT/FCI approach in an analogous way as described above for the CCSD(T)/FCI approach. We found that the former uncertain-

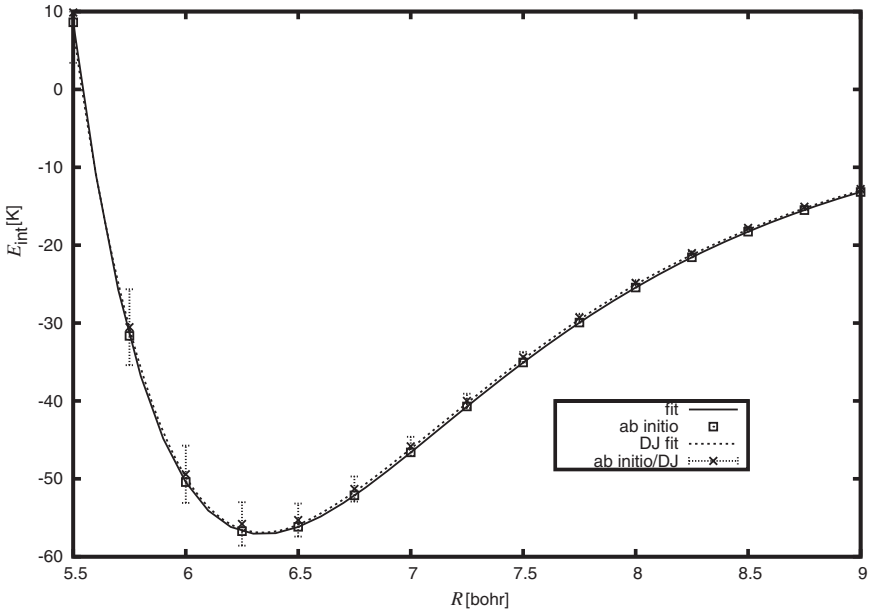

FIG. 3. Calculated and extrapolated (squares) and fitted (solid line) H_2-H_2 interaction energies for the angular configuration $(\theta_1,\theta_2,\phi)=(90,0,0)$ plotted as functions of R. The DJ potential, given in the Erratum of Ref. 22 (labeled "DJ fit"), and the DJ-level *ab initio* results and their uncertainties (obtained in the way described in the text and labeled by "x") are also drawn for comparison.

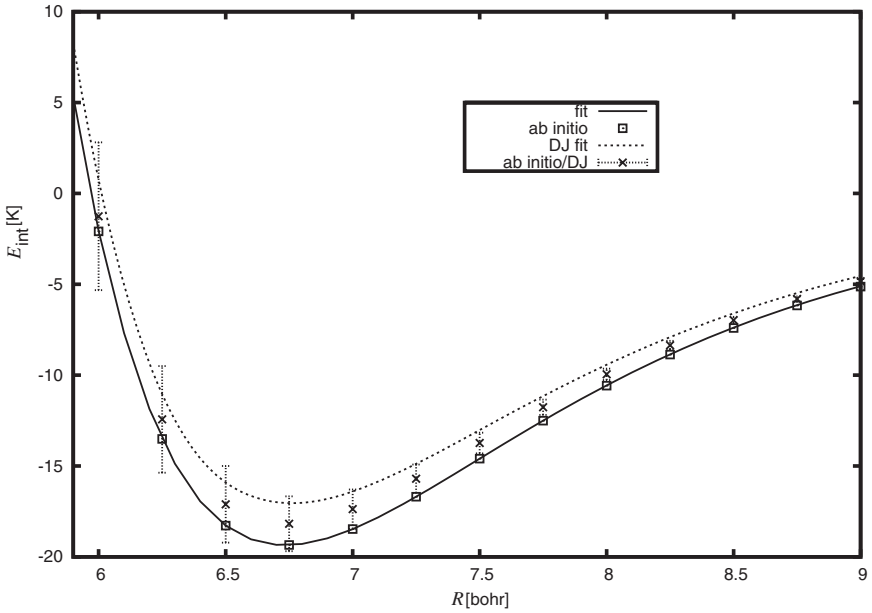

FIG. 4. Calculated and extrapolated (squares) and fitted (solid line) H_2-H_2 interaction energies for the angular configuration $(\theta_1,\theta_2,\phi)=(90,90,0)$ plotted as functions of R. The DJ potential, given in the Erratum of Ref. 22 (labeled "DJ fit"), and the DJ-level *ab initio* results and their uncertainties (obtained in the way described in the text and labeled by "x") are also drawn for comparison.

ties are larger than the latter in 25 out of 28 cases, indicating that, similarly as for the helium dimer,[12] the FCI correction beyond SAPT is harder to extrapolate to the CBS limit than an analogous FCI correction beyond CCSD(T). Therefore, in our potential fit we employed results from the CCSD(T)/FCI approach for all geometries, including the large-R ones. The SAPT calculations still form a valuable independent check of the accuracy of the *ab initio* points. One should emphasize that the differences between the SAPT/FCI and CCSD(T)/FCI results are very small: if SAPT/FCI results were plotted in Figs. 2–6, the curves would be indistinguishable from the CCSD(T)/FCI curves.

Without the FCI correction, the SAPT results for the H_2 dimer are generally of similar quality as the CCSD(T) results, but the van der Waals well predicted by SAPT is for most configurations too deep rather than too shallow, as is the case for CCSD(T). This is clearly visible in Figs. 7–11, where we plotted one-dimensional cuts through the potential analogous to Figs. 2–6. The plots contain both SAPT and CCSD(T) results calculated in the aug-cc-pV5Z+(33221) basis set. Note that these results are neither extrapolated nor augmented by the FCI correction. These plots show that the SAPT results are in most cases slightly farther from the benchmark than the CCSD(T) ones. When the intermonomer distance increases, the SAPT results become more accurate than the CCSD(T) ones. At $R=8.0$ bohrs, the accuracy of the CCSD(T) and SAPT theories is comparable: the absolute value of the FCI correction is smaller for CCSD(T) for 16 angular configurations from our grid and smaller for SAPT for the 12 remaining ones.

The SAPT calculations also provide important insights into the nature of the H_2-H_2 interaction as the SAPT interaction energy is a sum of well-defined terms corresponding to the electrostatic, induction, dispersion, and exchange effects.[56,57] These four physical components of the H_2-H_2 interaction potential, obtained from calculations in the

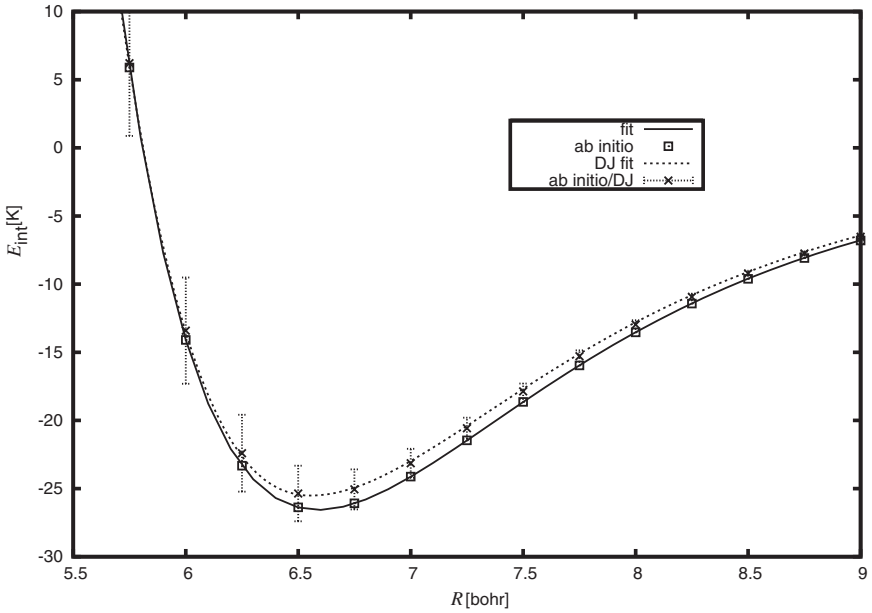

FIG. 5. Calculated and extrapolated (squares) and fitted (solid line) H_2-H_2 interaction energies for the angular configuration $(\theta_1,\theta_2,\phi)=(90,90,90)$ plotted as functions of R. The DJ potential, given in the Erratum of Ref. 22 (labeled "DJ fit"), and the DJ-level *ab initio* results and their uncertainties (obtained in the way described in the text and labeled by "x") are also drawn for comparison.

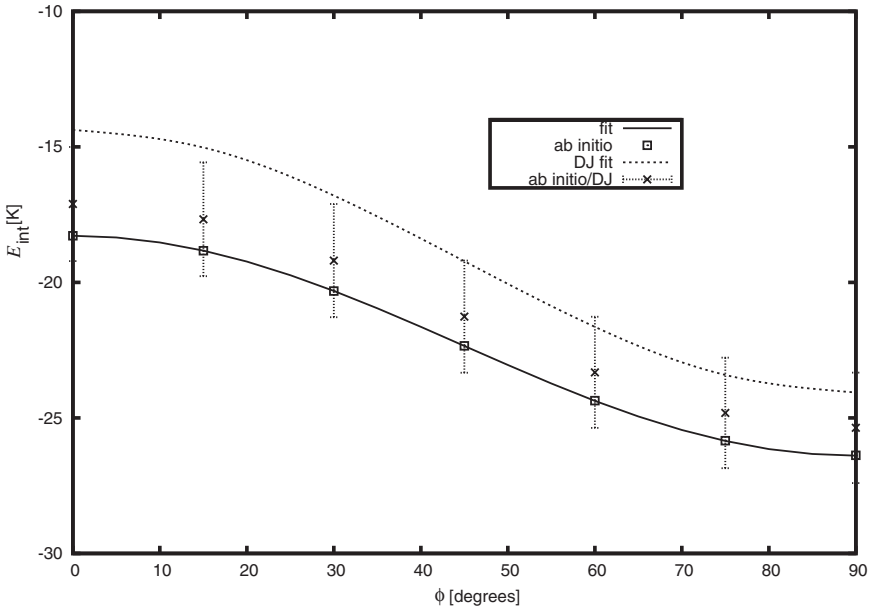

FIG. 6. Calculated and extrapolated (squares) and fitted (solid line) H_2-H_2 interaction energies for $R=6.5$ bohrs and $\theta_1=\theta_2=90$ plotted as functions of ϕ. The DJ potential, given in the Erratum of Ref. 22 (labeled "DJ fit"), and the DJ-level *ab initio* results and their uncertainties (obtained in the way described in the text and labeled by "x") are also drawn for comparison.

aug-cc-pV5Z+(33221) basis set, have been presented in Figs. 7–11 for the same one-dimensional cuts through the potential surface as depicted in Figs. 2–6. As expected, the main binding force in the H_2 dimer originates from the dispersion interactions, while the electrostatic and especially induction effects are much smaller. The electrostatic energy, exhibiting different signs for different relative orientations of the monomers, is the primary source of anisotropy of the potential. For asymptotic separations, this energy can be well described by the interaction of monomers' quadrupole moments. For finite R, this is largely also the case, as indicated by the potential cut along the ϕ coordinate, Figs. 6 and 11. The fitted potential of Diep and Johnson,[22] whose dependence on ϕ is assumed solely as a quadrupole-quadrupole term, recovers the anisotropy of the potential along this cut fairly well.

III. EXPLICITLY CORRELATED CALCULATIONS

In addition to the CCSD(T)/FCI and SAPT calculations, at the two geometries presented in Tables I and II, we used the method of explicitly correlated Gaussian (ECG) functions to better assess the achieved accuracy and as an independent check of our results. In the ECG method, the eigenfunction of an N-electron electronic Hamiltonian $\hat{H}$ is approximated as

$$\Psi = \mathcal{A}_N \left\{ \Xi^{N,S,M_S} \hat{P} \left[c_0 \phi_0 + \sum_{k=1}^{M} c_k \phi_k(1,2,\ldots N) \right] \right\}, \quad (4)$$

where $\mathcal{A}_N$ is the N-electron antisymmetrizer, Ξ^{N,S,M_S} is one of the N-electron spin functions corresponding to spin quantum numbers S and M_S, $\hat{P}$ is the point-group symmetry projector, and c_k are variational parameters. The functions $\phi_k(1,2,\ldots,N)$ have the form

$$\phi_k(1,2,\ldots,N) = \prod_{i=1}^{N} e^{-\alpha_{ki}|\mathbf{r}_i - \mathbf{A}_{ki}|^2} \prod_{i>j=1}^{N} e^{-\beta_{kij}|\mathbf{r}_i - \mathbf{r}_j|^2}, \quad (5)$$

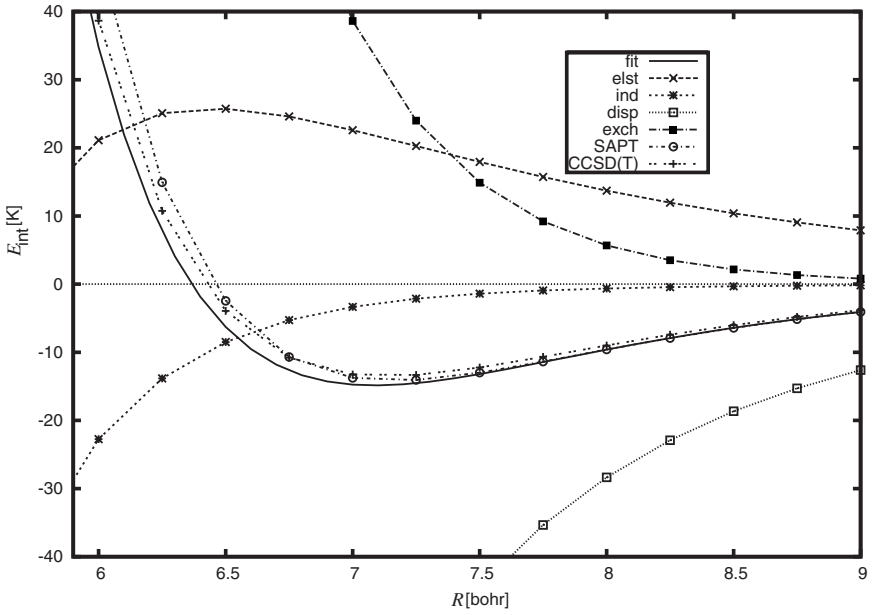

FIG. 7. Electrostatic, induction, dispersion, and exchange components of the SAPT interaction energy, their sum, and CCSD(T) interaction energy [all calculated in the aug-cc-pV5Z +(33221) basis set] for the angular configuration $(\theta_1,\theta_2,\phi)=(0,0,0)$ plotted as functions of R. Our total fitted potential is also drawn for comparison.

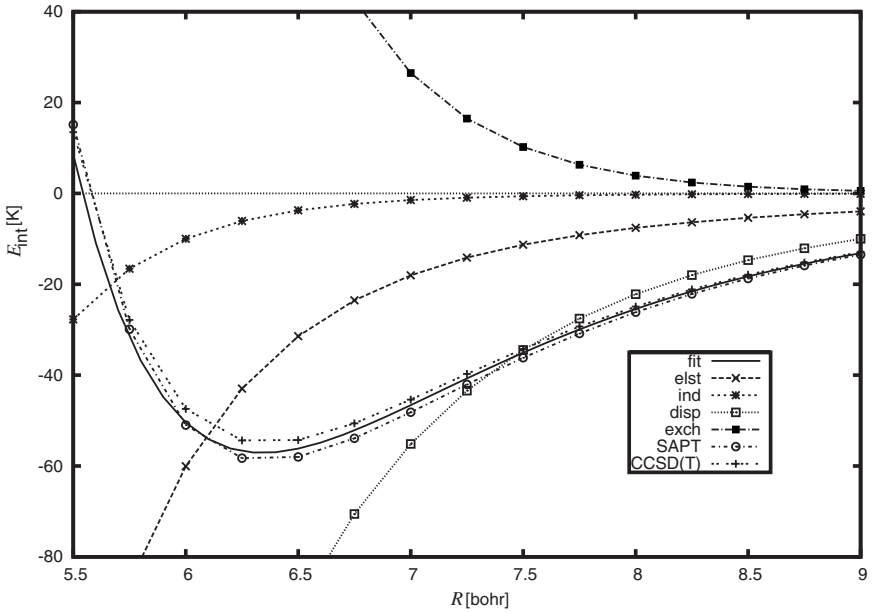

FIG. 8. Electrostatic, induction, dispersion, and exchange components of the SAPT interaction energy, their sum, and CCSD(T) interaction energy [all calculated in the aug-cc-pV5Z +(33221) basis set] for the angular configuration $(\theta_1,\theta_2,\phi)=(90,0,0)$ plotted as functions of R. Our total fitted potential is also drawn for comparison.

where α_{ki}, β_{kij}, and $\boldsymbol{A}_{ki}$ are nonlinear variational parameters. Due to the nonorthogonality of the basis functions and the necessity of a nonlinear optimization, this approach is too expensive for all but very small systems. However, it has been used with much success for a number of two-, three-, and four-electron systems (see Ref. 60 and references therein), where it invariably leads to accuracy levels unreachable for traditional orbital-based methods. The term ϕ_0 in Eq. (4), which does not undergo nonlinear optimization, can be included if a good approximation (also in the form of an ECG expansion) to the total wave function is available, which usually results in a faster convergence toward the $M=\infty$ limit. Such a favorable situation occurs in the case of weakly bound dimers, and it was exploited in Refs. 61 and 62 to construct very accurate wave functions of He_2. Specifically, using a two-electron ECG function of monomer X (where $X=$A or $X=$B),

$$\Psi_X = \mathcal{A}_2\left\{[\alpha(1)\beta(2)-\beta(1)\alpha(2)]\hat{P}\sum_{l=1}^{L} d_l\phi_l^X(1,2)\right\}, \quad (6)$$

where $\phi_l^X(1,2)$ are defined by Eq. (5) for $N=2$, the function ϕ_0 in Eq. (4) can be written as

$$\phi_0 = \sum_{l=1}^{L}\sum_{m=1}^{L} d_l d_m \phi_l^A(1,2)\phi_m^B(3,4). \quad (7)$$

Two-electron ECG optimizations are relatively inexpensive, and this so-called monomer-contraction (MC) technique allows the four-electron term ϕ_0 to be obtained with little effort.

Since L^2 is usually much larger than M, the calculation of the matrix elements $\langle\phi_0|\hat{H}|\phi_k\rangle$ with $k\neq 0$ (repeated many times during the nonlinear optimization of ϕ_k), scaling as L^2M, dominates the computational cost. Hence, it is essential

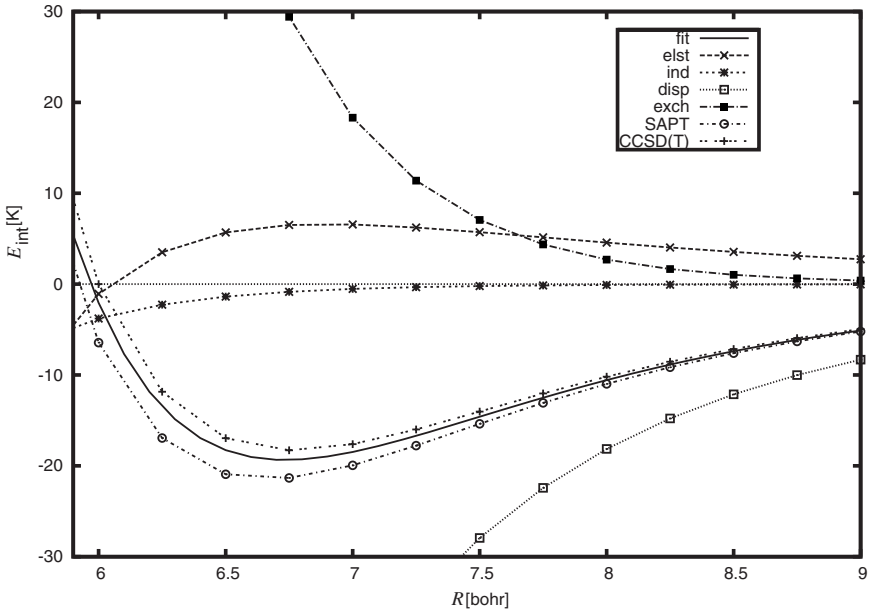

FIG. 9. Electrostatic, induction, dispersion, and exchange components of the SAPT interaction energy, their sum, and CCSD(T) interaction energy [all calculated in the aug-cc-pV5Z +(33221) basis set] for the angular configuration $(\theta_1,\theta_2,\phi)=(90,90,0)$ plotted as functions of R. Our total fitted potential is also drawn for comparison.

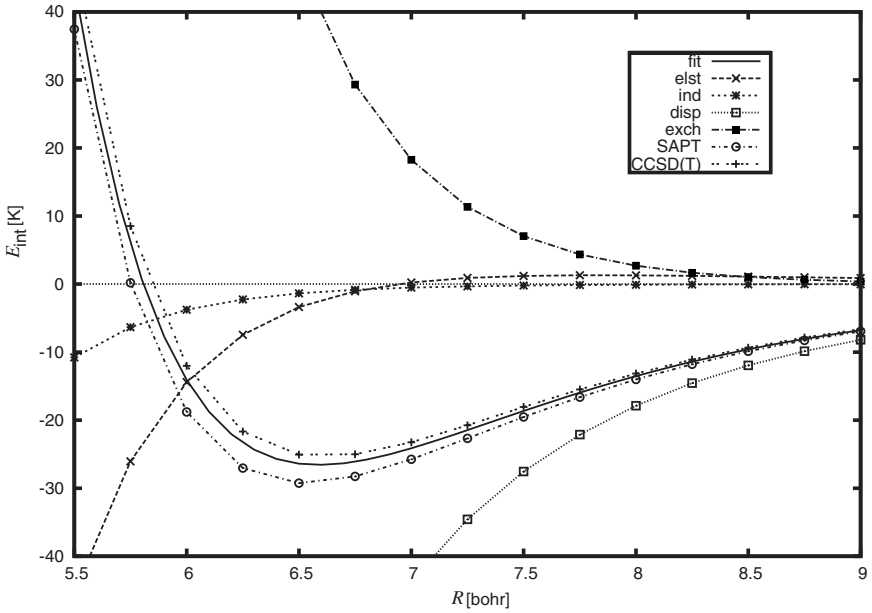

FIG. 10. Electrostatic, induction, dispersion, and exchange components of the SAPT interaction energy, their sum, and CCSD(T) interaction energy [all calculated in the aug-cc-pV5Z +(33221) basis set] for the angular configuration $(\theta_1, \theta_2, \phi)=(90,90,90)$ plotted as functions of R. Our total fitted potential is also drawn for comparison.

to generate accurate yet compact (very well optimized) monomer expansions of Eq. (6). Another possibility, which we explored for the first time in the present paper, is to perform a separate optimization of the contraction part of Eq. (4), i.e., of the function

$$\tilde{\phi}_0 = \mathcal{A}_4\{\Xi^{4,S,M_S}\hat{P}\phi_0\}, \tag{8}$$

using a more general form of Eq. (7):

$$\phi_0 = \sum_{l=1}^{L'} d_l \bar{\phi}_l^{\mathrm{A}}(1,2)\bar{\bar{\phi}}_l^{\mathrm{B}}(3,4), \tag{9}$$

where the functions $\bar{\phi}_l^{\mathrm{A}}$ and $\bar{\bar{\phi}}_l^{\mathrm{B}}$ still have the form of Eq. (5) but are mutually independent and Eq. (9) does not represent a product of two two-electron expansions. The linear parameters d_l and the nonlinear parameters contained in the function of Eq. (9) are optimized by minimizing the sum of monomer energies, $\langle\tilde{\phi}_0|\hat{H}^0|\tilde{\phi}_0\rangle/\langle\tilde{\phi}_0|\tilde{\phi}_0\rangle$, where $\hat{H}^0$ is the Hamiltonian of two noninteracting monomers,

$$\hat{H}^0 = \hat{H}^{\mathrm{A}} + \hat{H}^{\mathrm{B}}. \tag{10}$$

During the optimization, in order to avoid the exchange interactions between A and B, the permutations mixing electrons from both monomers are switched off in the antisymmetrizer, i.e., $\mathcal{A}_4$ is replaced by

$$\tilde{\mathcal{A}}_4 = \tfrac{1}{8}(1-P_{12})(1-P_{34})(1+P_{13}P_{24}), \tag{11}$$

where P_{ij} interchanges electrons i and j.

The use of the optimized form of function (8) leads to much shorter expansions of the monomer contraction at the same level of accuracy, as discussed below. Moreover, separating the optimization of $\tilde{\phi}_0$ from the optimization of the other terms in Eq. (4) is advantageous for at least two reasons. First, the optimization of $\tilde{\phi}_0$ is computationally much less demanding than that of the remaining part of Ψ because

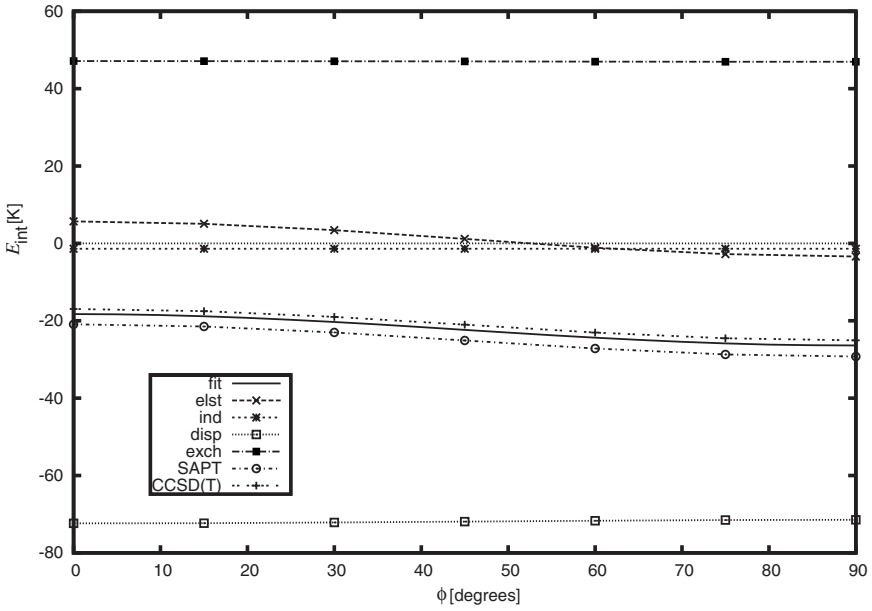

FIG. 11. Electrostatic, induction, dispersion, and exchange components of the SAPT interaction energy, their sum, and CCSD(T) interaction energy [all calculated in the aug-cc-pV5Z +(33221) basis set] for R=6.5 bohrs and $\theta_1=\theta_2=90$ plotted as functions of ϕ. Our total fitted potential is also drawn for comparison.

TABLE IV. Comparison of ECG expansion lengths and errors δE (in kelvins, with respect to the exact energy of two noninteracting H_2 molecules) obtained by optimizing the H_2 wave function (upper) and by optimizing the antisymmetrized wave function of two noninteracting H_2 monomers (lower). The intramonomer distance was 1.448 736 bohrs.

Ψ_{H_2} L terms	$\Psi_{H_2}\cdot\Psi_{H_2}$ $4L^2$ terms[a]	δE
75	22 500	0.753
106	44 944	0.150
150	90 000	0.029
212	179 776	0.0065
300	360 000	0.0009
$\Psi_{H_2+H_2}$ L' terms	$2L'$ terms[b]	δE
1200	2400	0.523
2400	4800	0.107
3600	7200	0.047

[a]Terms of the form of Eq. (7), with relaxed point-group symmetry.
[b]Terms of the form of Eq. (9), with relaxed point-group symmetry.

of many vanishing terms in the integrals (due to the lack of interactions between monomers A and B). Second, $\tilde{\phi}_0$ is optimized only once and then used for each geometry of the dimer.

Table IV compares the errors of the sum of the monomer energies (with respect to the exact energy of two isolated H_2 molecules) in the older MC method and in the new approach. The energy of the hydrogen molecule with the interatomic distance equal to 1.448 736 bohrs was calculated by us from a carefully optimized 2400-term ECG expansion and amounts to −1.174 077 876 74 hartrees (1 hartree $\approx 4.359\,744\times10^{-18}$ J), where all the listed digits are exact. Note that H_4 has (in general) C_1 symmetry, while in the optimizations of the Ψ_X and $\tilde{\phi}_0$ functions the Gaussian centers were restricted to the molecular axis and the symmetry operator $\hat{P}=1+\hat{\imath}$ ($\hat{\imath}$ being the inversion through the molecular midpoint) was used to enforce the Σ_g^+ symmetry. Therefore, after switching off the symmetry, these functions have double expansion lengths ($2L$ and $2L'$, respectively). Table IV illustrates clearly that the new method of generating the monomer contraction, i.e., the use of $\tilde{\phi}_0$ from the direct optimization of the four-electron partly antisymmetrized wave function of noninteracting monomers, yields large computational gains. Namely, to achieve a similar accuracy of the noninteracting part of the dimer wave function, much shorter expansions (by about one order of magnitude) of the zeroth term in Eq. (4) are necessary compared to the old MC method. Thus, the optimization of the remaining terms, describing the intermonomer interaction, is faster by the same factor.

To optimize the interaction part, we used the 7200-term (in C_1 symmetry) function $\tilde{\phi}_0$, yielding the energy of noninteracting monomers with an error of 47 mK, and expansions of different lengths for the interaction part, with $M=150$, 300, 600, 1200, and 2400. In the final step, the 7200-term function $\tilde{\phi}_0$ was replaced by a more accurate, 360 000-term function resulting from a product of two 300-term monomer functions and exhibiting the error of only 0.9 mK (see Table IV). The dimer energies became lower by 38 (for the linear configuration and $M=2400$) to 44 mK. The fact that this lowering was smaller than 46 mK (the difference between the accuracies of the two monomer contractions) results from partial recovery of the missing monomer effects during the dimer optimization, especially for longer interaction part expansions. Table V presents the final results for the same T-shaped and linear configurations as in Tables I and II. The interaction energies were obtained by subtracting twice the *exact* energy of H_2 quoted above. Hence, all the results for finite values of M given in Table V represent strict upper bounds to the respective exact values.

TABLE V. Interaction energies (in kelvins) for the T-shaped configuration with the COM distance of 6.4 bohrs and for the linear configuration with the COM distance of 6.0 bohrs obtained using the product of two 300-term H_2 wave functions (reproducing the sum of the monomer energies with an error of 0.9 mK) and M terms in the interaction part of Eq. (4). The ECG bases were optimized using the 7200-term $\tilde{\phi}_0$. Boldface numbers denote the recommended ECG values. The orbital CCSD(T)/FCI results are included for comparison. The intramonomer distance was 1.448 736 bohrs.

M	T-shaped	Linear
150	−54.3468	37.0560
300	−55.7869	35.6170
600	−56.4469	35.0862
1200	−56.7062	34.8889
2400	−56.8213	34.8286
Extrapolated	-56.895 ± 0.047	34.790 ± 0.025
Corrected for optimization incompleteness	$\mathbf{-56.905\pm0.057}$ [a]	$\mathbf{34.785\pm0.030}$ [b]
Orbital (whole-grid basis)	-56.96 ± 0.16 [c]	34.74 ± 0.30 [c]
Orbital (largest basis)	-56.93 ± 0.19 [c]	34.77 ± 0.32 [c]

[a]With an optimization incompleteness correction of -0.010 ± 0.010 (see text).
[b]With an optimization incompleteness correction of -0.005 ± 0.005 (see text).
[c]Estimated as $\sigma_{CCSD(T)}+2\sigma_{FCI}$, where $\sigma_{CCSD(T)/FCI}$ is the magnitude of the difference between the result computed in a basis set with cardinal number X and the results extrapolated from bases with cardinal numbers $X-1$ and X.

There is some empirical evidence[63] that for optimized ECG functions, the ratio

$$\eta_M = \frac{E_{M/4} - E_{M/2}}{E_{M/2} - E_M} \tag{12}$$

is approximately independent of M. Thus, the increments in energies upon doubling of M approximately form a geometric series with the quotient $1/\eta_M$. In the present case, the six values of η_M that can be derived from the data in Table V range from 2.18 to 3.27. The assumption of $\eta_M = 2.18$ for all $M > 2400$ leads to the extrapolated values of −56.9188 and 34.7775 K for the T-shaped and linear configurations, respectively, whereas taking $\eta_M = 3.27$ leads to −56.8720 and 34.8020 K. At each configuration, we take the midpoint of the two predictions as the recommended result of our CBS extrapolation scheme and the full difference between these predictions as the uncertainty estimate.

The CBS extrapolation described above would have been final if our optimizations of nonlinear parameters had been completely converged for each M. This is not the case, however, and we have to take into account small residual uncertainties, especially in the case of the very time-consuming optimizations of the $M = 2400$ functions. By analyzing the optimization cycle convergence, we estimated these residual uncertainties to be about 5–15 mK for the T-shaped configuration and several millikelvins for the linear geometry. With some safety margin, we assumed that completely converged optimizations could lower the energy by any value between 0 and 20 mK for the T-shaped structure and by any value between 0 and 10 mK for the linear one. Therefore, we have lowered the CBS limits determined above by 0.010 and 0.005 K for the T-shaped and linear configurations, respectively, and enlarged the respective error bars by 0.010 and 0.005 K. This leads to the recommended CBS limits of -56.905 ± 0.057 and 34.785 ± 0.030 K for the two geometries. The very small residual error in the monomer contraction used in the final calculations [not larger than 0.9 mK, see the discussion before Eq. (12)] can be neglected relative to the much larger uncertainties discussed so far.

As can be seen in Table V, the ECG calculations are significantly more accurate than the orbital calculations, with uncertainties three times smaller for the T-shaped geometry and ten times smaller for the linear geometry. The recommended ECG values are higher than the orbital predictions. It is seen in Tables I and II that, at least for the basis sets with bond functions, the CCSD(T) energies converge from above while the FCI corrections converge from below and that this is true for both calculated and extrapolated results. Hence, even though the FCI correction is about 30 times (10 times) smaller in absolute value than the CCSD(T) energy for the T-shaped (linear) geometry, the former quantity seems worse converged. One reason is undoubtedly the fact that smaller basis sets were used to evaluate the FCI part. As already mentioned in Sec. II, it is also possible that the convergence of this quantity is inherently slower and that the X^{-3} extrapolation used by us underestimates the distances between the calculated values and the CBS limits. On the other hand, the predictions of both methods in Table V are mutually consistent; in fact, the orbital results in the largest basis set are within the ECG error bars and those in the whole-grid basis are nearly so. This consistency indicates that our method of extrapolating the orbital results and estimating their uncertainties is indeed reliable but probably too conservative as the magnitude of the difference between the ECG and the whole-grid-level orbital results is three (six) times smaller than the estimated uncertainty in the latter result for the T-shaped (linear) configuration.

IV. FITTING PROCEDURE

To develop a functional representation of our potential, we used the set of 213 *ab initio* data points described in Sec. II. The functional form of the fit, the same as used in an earlier work on the CO_2 dimer,[64] is composed of a short-range part, U_{sh}, and an asymptotic part, U_{as},

$$U(R, \theta_1, \theta_2, \phi) = U_{sh}(R, \theta_1, \theta_2, \phi) + U_{as}(R, \theta_1, \theta_2, \phi). \tag{13}$$

The short-range part is defined by

$$U_{sh}(R, \theta_1, \theta_2, \phi) = G(R, \theta_1, \theta_2, \phi) e^{D(\theta_1, \theta_2, \phi) - B(\theta_1, \theta_2, \phi)R}, \tag{14}$$

where G is a polynomial in R with orientation-dependent coefficients,

$$G(R, \theta_1, \theta_2, \phi) = \sum_{i=0}^{3} R^i \sum_{l_1 \geq l_2, l} g_i^{l_1 l_2 l} A^{s}_{l_1 l_2 l}(\theta_1, \theta_2, \phi). \tag{15}$$

The angular functions $A_{l_1 l_2 l}(\theta_1, \theta_2, \phi)$, members of a complete basis set, are defined as

$$A_{l_1 l_2 l}(\theta_1, \theta_2, \phi) = \sum_{m=-l_<}^{l_<} \begin{pmatrix} l_1 & l_2 & l \\ m & -m & 0 \end{pmatrix} C_{l_1 m}(\theta_1, \phi) \times C_{l_2 -m}(\theta_2, 0) \tag{16}$$

and are symmetrized according to

$$A^{s}_{l_1 l_2 l}(\theta_1, \theta_2, \phi) = A_{l_1 l_2 l}(\theta_1, \theta_2, \phi) + A_{l_1 l_2 l}(\theta_2, \theta_1, \phi), \tag{17}$$

where $C_{lm}(\theta, \phi) = (4\pi)^{1/2}(2l+1)^{-1/2} Y_{lm}(\theta, \phi)$ is the spherical harmonic in the Racah normalization, $l_<$ is the smaller of l_1 and l_2, and the large brackets denote the Wigner $3j$ symbol. The exponential parameters of Eq. (14) are expanded as

$$D(\theta_1,\theta_2,\phi)=\sum_{l_1\geq l_2,l} d^{l_1l_2l}A^{s}_{l_1l_2l}(\theta_1,\theta_2,\phi), \tag{18}$$

$$B(\theta_1,\theta_2,\phi)=\sum_{l_1\geq l_2,l} b^{l_1l_2l}A^{s}_{l_1l_2l}(\theta_1,\theta_2,\phi). \tag{19}$$

The asymptotic part of the potential is defined as

$$U_{\mathrm{as}}(R,\theta_1,\theta_2,\phi)=\sum_{l_1l_2,l,n} f_n[B(\theta_1,\theta_2,\phi)R]\frac{C_n^{l_1l_2l}}{R^n} \times A_{l_1l_2l}(\theta_1,\theta_2,\phi), \tag{20}$$

where $f_n[x]$ is the Tang–Toennies damping factor,[65]

$$f_n[x]=1-e^{-x}\sum_{k=0}^{n}\frac{x^k}{k!}. \tag{21}$$

We have used the nonsymmetrized functions $A_{l_1l_2l}(\theta_1,\theta_2,\phi)$ in Eq. (20) since the literature van der Waals constants $C_n^{l_1l_2l}$ utilized by us (see below) are defined relative to such functions. The values of n ranged from 5 to 10 in Eq. (20). The constants $C_n^{l_1l_2l}$ were held fixed during the fitting process. In the angular expansion of Eqs. (15), (18), and (19), we used eight, five, and five functions $A^{s}_{l_1l_2l}$, respectively. The total number of adjustable fit parameters amounts to 42 (8 $g_i^{l_1l_2l}$ parameters for each of the four powers of R, 5 $d^{l_1l_2l}$ parameters, and 5 $b^{l_1l_2l}$ parameters). The fitting algorithm, including the choice of the initial values of the adjustable parameters, was the same as that employed in Ref. 64. The *ab initio* interaction energies $E(R,\theta_1,\theta_2,\phi)$ were given weights $w=w_Ew_R$, with the energy-dependent and distance-dependent weighting factors defined as

$$w_E=\begin{cases}\left(\dfrac{E_w}{E}\right)^2 & \text{for } E>E_w\\ e^{2(1-E/E_w)} & \text{for } E\leq E_w,\end{cases} \tag{22}$$

$$w_R=\left(\frac{R}{R_w}\right)^3, \tag{23}$$

where $E_w=200$ K and $R_w=6.5$ bohrs.

The van der Waals constants $C_n^{l_1l_2l}$ were taken from Table IX of Ref. 66 (note that the coefficients $V^n_{l_al_bl}$ from that table must be multiplied by the factor $\sqrt{2l+1}$ to yield $C_n^{l_1l_2l}$) except for the constants with $n=6$ and those originating from the interaction of the permanent multipole moments of the monomers. Note that these constants describe all asymptotic interactions: electrostatic, induction, and dispersion. We used newer, more accurate values of C_6 computed by Bishop and Pipin,[67] and the values of C_5, C_7, and C_9 were derived from the multipole moments calculated by Komasa and Thakkar.[68] In both cases, the authors employed highly accurate expansions in explicitly correlated wave functions and these are the highest-level asymptotic data available. All the van der Waals constants used in the fit were calculated at the intramonomer distance equal to 1.449 bohrs rather than 1.448 736 bohrs as employed in our *ab initio* calculations. We checked that the uncertainties introduced in this way are negligible by comparing two sets of *ab initio* energies at $R=12$ bohrs using both monomer geometries.

To check how the chosen asymptotic expansion agrees with our *ab initio* interaction energies, we calculated ten additional *ab initio* points for $R=12$ and 15 bohrs and (θ_1,θ_2,ϕ) equal to (0,0,0), (90,0,0), (90,90,0), (90,90,90), and (60,30,180), applying exactly the same method as in the calculations for all the grid points used in the fit. The differences between the energies calculated *ab initio* and from the asymptotic formula are 1% or less, with the exception of the (0,0,0) configurations, where these differences are larger: 3.5% (2.4 mK) at $R=12$ bohrs and 1.3% (1.6 mK) at $R=15$ bohrs. The reason the (0,0,0) configuration has the largest relative deviations is most likely the fact that for a given R the closest distance between atoms belonging to different monomers is smallest in this case, and that the electrostatic contribution is repulsive for this configuration. Moreover, for the (0,0,0) angular orientation the sign of the interaction energy changes from negative to positive between $R=12$ and 15 bohrs, and the magnitude of the interaction energy at $R=12$ bohrs is very small compared to those in other angular configurations tested. To improve the fit for larger R, one would have to compute more accurately the asymptotic constants. In particular, one needs the C_{11} and C_{12} constants since the C_{10}/R^{10} term contributes as much as −45 mK at $R=12$ bohrs in the (0,0,0) configuration. These improvements would have a small effect, however, on our overall fit since at R smaller than about 10 bohrs the short-range part of the fit makes up for all inaccuracies of the asymptotic constants.

The value of the root-mean-square error (rmse) of our fit for all the fitted points amounts to 1.64 K, while the rmse for all the points with negative interaction energies is 0.018 K. For most of the points, the difference between the calculated energy and the fit value is less than the uncertainty in the former. The same is true for the ten test configurations $\boldsymbol{G}_1,\ldots,\boldsymbol{G}_{10}$, chosen as described in Sec. II, which have the rmse of 0.494 K for all points and 0.016 K for points with negative energies. It appears, therefore, that the quality of our fit is more than adequate despite the relatively small number of grid points computed by us. The high quality of the fit is also visible in Figs. 2–6, where the calculated energy points lie virtually exactly on the fit curve. The global minimum of the fit is −57.118 K for $R=6.34$ bohrs and the angular configuration (90,0,0).

In addition to the four-dimensional interaction energy surface $U(R,\theta_1,\theta_2,\phi)$ of H_2–H_2, we have developed an isotropic potential for this complex. It should be noted that the isotropic potential cannot be obtained directly from $U(R,\theta_1,\theta_2,\phi)$ by setting all anisotropic coefficients equal to 0. This is because the analytical form of the surface, defined by Eqs. (13)–(21), is a nonlinear function of the angular basis set components. Therefore, to obtain the isotropic po-

tential $U_{\rm iso}(R)$, the following procedure has been employed. For values of R_i every 0.25 bohr between 4.0 and 20.0 bohrs, the potential surface $U(R_i,\theta_1,\theta_2,\phi)$ was numerically integrated over the angular coordinates with the isotropic basis function $A_{000}^{\rm s}\equiv 2$ to obtain values of the isotropic interaction energy, $U_{\rm iso}(R_i)$. To analytically represent this set of isotropic interaction energies, we assumed a formula of the same form as that used to represent $U(R,\theta_1,\theta_2,\phi)$ but with terms other than those containing $A_{000}^{\rm s}$ omitted. The long-range component of such an isotropic potential is defined by the coefficients C_n^{000}, $n=6,8,10$. The short-range part has six free parameters whose values have been fitted to $U_{\rm iso}(R_i)$.

The FORTRAN subroutines computing our $U(R,\theta_1,\theta_2,\phi)$ and $U_{\rm iso}(R)$ potentials are included in the supporting information.[54]

V. SECOND VIRIAL COEFFICIENT

A. Background

One way to test a pair potential is to compare with experiment its prediction of the second virial coefficient $B(T)$, a thermodynamic quantity (representing the leading-order deviation from the ideal-gas law) that depends only on the pair potential. For sufficiently accurate potentials, it is possible to reverse the typical process and use pair potentials to predict values of $B(T)$ with better accuracy than obtained by experiment.[10,69] Fairly good $B(T)$ data for pure hydrogen exist from a number of sources over a range of temperatures; here we will compare with the two studies generally considered to be the most accurate and comprehensive. Goodwin *et al.*[70] analyzed previously published compressibility data for *para*-hydrogen[71] and reported second virial coefficients at temperatures ranging from 24 to 100 K. Michels *et al.*[72] reported $B(T)$ for normal hydrogen between approximately 98 and 423 K.

The calculation of $B(T)$ for hydrogen from a pair potential is complicated by quantum effects due both to the small mass and small moment of inertia of H_2. While $B(T)$ can be calculated at the fully quantum level for atomic gases, such calculations would be difficult for anisotropic potentials. Harvey and Hodges[25] used a semiclassical approach to analyze the effect of anisotropy on $B(T)$ for the DJ pair potential. Significant deviations were observed between the calculated second virial coefficients and the experimental data of Michels *et al.*,[72] but it was unclear whether this was due to incomplete inclusion of quantum effects or due to deficiencies in the DJ potential. The DJ paper on the H_2–H_2 potential[22] also included PIMC calculations of $B(T)$, although restricted to isotropic potentials. Since that work, Schenter[31] showed how to calculate $B(T)$ for anisotropic potentials with the PIMC method, which should approximate the exact quantum solution in the limit of infinite computer time. Here, we use the PIMC technique to produce values of $B(T)$ that are accurate to within small uncertainties. In order to elucidate the role of quantum effects and anisotropy, we also report semiclassical calculations with our pair potential and fully quantum calculations with the isotropic version of our potential.

B. Path-integral methodology

The fully quantum expression of the second virial coefficient for a gas of rigid linear rotors of mass m and inertia moment I in the approximation where the Bose or Fermi nature of the particles can be neglected is given by[31,73]

$$B(T)=-\frac{1}{2}\left[\frac{\Lambda_\mu^3}{q_{\rm rot}^2}{\rm Tr}(\exp(-\beta\hat{h}_2)-1)\right], \tag{24}$$

where $\beta=1/k_{\rm B}T$, $\Lambda_\mu=h/\sqrt{2\pi\mu k_{\rm B}T}$ is the de Broglie thermal wavelength of a particle with reduced mass $\mu=m/2$, and $q_{\rm rot}=\Sigma_l(2l+1)\exp[-\beta\Upsilon l(l+1)]$ is the partition function of a free rotor given the rotational constant $\Upsilon=\hbar^2/2I$. The Hamiltonian $\hat{h}_2$ is given by

$$\hat{h}_2=\frac{\hat{\mathbf{p}}^2}{2\mu}+\frac{\hat{\mathbf{J}}_1^2}{2I}+\frac{\hat{\mathbf{J}}_2^2}{2I}+\hat{U}(r,\Omega_1,\Omega_2), \tag{25}$$

where $\hat{U}(r,\Omega_1,\Omega_2)$ is the intermolecular potential as a function of the COM distance r and the orientations Ω_1 and Ω_2 of the two rotors. The operators $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ describe the angular momenta of the rotors and $\hat{\mathbf{p}}$ is the linear momentum conjugated to the COM separation.

The trace appearing in Eq. (24) can be evaluated using path-integral methods and we chose to proceed by performing a Trotter factorization of order P and inserting completeness relations between the various terms. Using as basis the complete set given by the product of eigenfunctions of linear and angular positions $|\mathbf{x}^{(i)}\Omega_1^{(i)}\Omega_2^{(i)}\rangle$, where the index i denotes each of the P Trotter factors, the final expression for the second virial coefficient turns out to be

$$B(T)=-\frac{1}{2}\int d^3\mathbf{r}d^3\Delta^{(1)}d^3\Delta^{(2)}\cdots d^3\Delta^{(P-1)}d\Omega_1^{(1)}d\Omega_1^{(2)}\cdots d\Omega_1^{(P)}d\Omega_2^{(1)}d\Omega_2^{(2)}\cdots d\Omega_2^{(P)}$$
$$\times F(\mathbf{\Delta})\varrho(\mathbf{\Omega}_1)\varrho(\mathbf{\Omega}_2)\left(\exp\left[-\frac{\beta}{P}\sum_{k=1}^{P}U(|\mathbf{x}^{(k)}|,\Omega_1^{(k)},\Omega_2^{(k)})\right]-1\right), \tag{26}$$

where we have introduced the coordinates $\mathbf{\Delta}^{(i)}=\mathbf{x}^{(i+1)}-\mathbf{x}^{(i)}$ with the understanding that $\mathbf{x}^{(P+1)}=\mathbf{x}^{(1)}$ and $\Omega^{(P+1)}=\Omega^{(1)}$. The variable $\mathbf{r}\equiv\mathbf{x}^{(1)}$ has been introduced for later convenience and we have denoted by $\mathbf{\Delta}$, $\mathbf{\Omega}_1$, and $\mathbf{\Omega}_2$ the set of all the coordinates $\mathbf{\Delta}^{(i)}$, $\Omega_1^{(i)}$, and $\Omega_2^{(i)}$, respectively. Notice that one can write $\mathbf{x}^{(j)}=\mathbf{r}+\sum_{n=1}^{j-1}\mathbf{\Delta}^{(n)}$.

The quantities F and ϱ are expressed as[74,75]

$$F(\mathbf{\Delta})=\Lambda_\mu^3\prod_{i=1}^{P}\langle\mathbf{x}^{(i)}|\exp\left(-\beta\frac{\hat{\mathbf{p}}^2}{2\mu P}\right)|\mathbf{x}^{(i+1)}\rangle$$

$$=\Lambda_\mu^3\left(\frac{P^{3/2}}{\Lambda_\mu^3}\right)^P\exp\left(-\frac{\pi P}{\Lambda_\mu^2}\sum_{i=1}^{P}|\mathbf{\Delta}^{(i)}|^2\right), \tag{27}$$

$$\varrho(\mathbf{\Omega})=\frac{1}{q_{\mathrm{rot}}}\prod_{i=1}^{P}\langle\Omega^{(i)}|\exp\left(-\beta\frac{\hat{\mathbf{J}}^2}{2IP}\right)|\Omega^{(i+1)}\rangle$$

$$=\frac{1}{q_{\mathrm{rot}}}\prod_{i=1}^{P}\sum_{j=0}^{\infty}\frac{2j+1}{4\pi}P_j(\cos\,\theta_{i,i+1})$$

$$\times\exp[-\beta j(j+1)\Upsilon/P], \tag{28}$$

where $P_j(\cos\,\theta)$ is a Legendre function and $\theta_{i,i+1}$ is the angle between the directions described by the solid angles $\Omega^{(i)}$ and $\Omega^{(i+1)}$. The functions F and ϱ describe the probability distribution of the internal coordinates of an ideal-gas ring polymer[74] and the probability density for the orientation of the rotors,[75] respectively.

The integration appearing in Eq. (26) is performed over a bundle of P rotors whose COMs are fixed at the origin and whose directions $\mathbf{\Omega}_1$ are distributed according to the function $\varrho(\mathbf{\Omega}_1)$. These rotors interact with a ring polymer of P beads whose relative distances $\mathbf{\Delta}$ are distributed according to the function $F(\mathbf{\Delta})$. Each of the beads of this ring polymer is associated with a rotor, and the orientations of these P rotors are distributed according to the function $\varrho(\mathbf{\Omega}_2)$. The interaction between these rotors is such that it only occurs if they have the same superscript (i) (corresponding to the same "time slice" in the path-integral formalism).

One can then define, similarly to what has been done by Schenter,[31] an effective radial potential $\bar{U}(r)$ such that

$$\exp[-\beta\bar{U}(|\mathbf{r}|)]=\left\langle\exp\left[-\frac{\beta}{P}\sum_{i=1}^{P}U(|\mathbf{x}^{(i)}|,\Omega_1^{(i)},\Omega_2^{(i)})\right]\right\rangle, \tag{29}$$

where the average is performed over the ring polymer conformations $\mathbf{\Delta}$ and the rotor orientations $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ according to the distribution functions F and ϱ. With this definition, Eq. (26) can be written as

$$B(T)=-2\pi\int_0^\infty r^2dr(\exp[-\beta\bar{U}(r)]-1), \tag{30}$$

which is equivalent to the expression given by Schenter.[31]

The main difference between our approach and the one of Diep and Johnson[22] and of Schenter[31] is that we have performed a canonical transformation to the COM and relative coordinates before applying the Trotter factorization [see Eq. (24)], following the idea first put forward by Fosdick and Jordan[76] in their pioneering calculation of the second virial coefficient of ^{4}He using path-integral methods. The COM coordinate can be integrated out because it commutes with all the other coordinates, and one is left with the relative coordinate $\mathbf{r}$ only, greatly reducing the computational effort needed to evaluate the second virial coefficient. Notice that in the path-integral formalism this relative coordinate describes the position of one of the beads of the ring polymer that corresponds to the relative distance between the two original quantum particles and not the distance between the centroids.

We have calculated the Boltzmann factor of the effective potential of Eq. (29) at 400 equispaced values of the radial distance r up to a cutoff of $r_{\max}=40$ Å and interpolated the values at arbitrary distances using a cubic spline. The second virial coefficient has been obtained by evaluating the integral in Eq. (30) with an adaptive Gauss–Kronrod method. For each value of the distance r, the average has been performed by using $N=10^5$ ring polymers and rotational configurations. The ideal-gas ring polymers have been generated using the interpolation formula due to Levy and reported in Ref. 76, whereas the rotational directions have been generated by performing a hybrid PIMC simulation on 64 free rotors, assuming that all the angular momentum states $\mathbf{J}$ could be reached.[77,78] The value of the Trotter index P was chosen at each temperature T according to $P=3600\ \mathrm{K}/T$. The statistical uncertainty in the quantum value of $B(T)$ has been evaluated by performing a radial integration of the standard error of the mean of the Boltzmann factor evaluated in Eq. (29).

C. Other calculation methods

We have also computed the semiclassical approximation to the second virial coefficient, adding to the classical $B(T)$ the leading quantum correction proportional to $\hbar^2$. We used the expressions reported in Ref. 73 together with the centrifugal term derived by Pack.[79]

In addition, in order to quantify the effect of anisotropy, we performed fully quantum calculations of $B(T)$ for the isotropic potential derived in Sec. IV. We used the phase-shift method described in Ref. 73; the details of the calculations are as described by Hurly and Mehl.[10]

D. Results

Table VI shows calculated values of $B(T)$ at several temperatures from the PIMC method. The listed uncertainties represent the uncertainty in the convergence of the PIMC calculations; possible systematic uncertainty due to errors in the pair potential is not included. The table also shows results from a first-order semiclassical perturbation calculation of quantum effects and from fully quantum calculations with the isotropic potential.

From these calculations, we are in a position to address the question that Harvey and Hodges[25] were unable to answer; namely, whether the disagreement with experiment of their semiclassical calculation of $B(T)$ from the DJ potential

TABLE VI. Values of the second virial coefficient $B(T)$ for H_2 obtained with the PIMC method, with the semiclassical approximation, and with a fully quantum calculation employing the isotropic approximation to the potential.

Temperature (K)	$B(T)$ (PIMC) (cm^3/mol)	$B(T)$ (semiclassical) (cm^3/mol)	$B(T)$ (quantum, isotropic) (cm^3/mol)
15		1235.4	−221.7
30		−16.42	−80.32
50	−33.43 ± 0.13	−23.73	−33.20
75	−12.08 ± 0.12	−9.61	−11.91
100	−2.15 ± 0.07	−1.18	−2.02
125	3.47 ± 0.06	3.92	3.56
150	6.98 ± 0.05	7.24	7.07
175	9.34 ± 0.04	9.52	9.44
200	11.02 ± 0.04	11.15	11.12
225	12.26 ± 0.04	12.35	12.34
250	13.20 ± 0.03	13.25	13.27
273.16	13.85 ± 0.03	13.90	13.92
300	14.45 ± 0.03	14.49	14.52
350	15.21 ± 0.03	15.25	15.28
400	15.70 ± 0.02	15.72	15.77
450	16.00 ± 0.02	16.03	16.07
500	16.19 ± 0.02	16.21	16.25
550	16.29 ± 0.02	16.31	16.35
600	16.34 ± 0.02	16.36	16.40

was due to deficiencies in the DJ potential, due to neglect of higher-order quantum effects, or due to both.

Figure 12 is plotted analogously to Fig. 1 in the paper of Harvey and Hodges.[25] In particular, the zero line in Fig. 12 represents the semiclassical calculation with the DJ potential.[22] There are two main differences. The experimental data from Michels *et al.*[72] are now shown with error bars corresponding to a standard uncertainty with coverage factor $k=2$ (95% confidence interval), whereas Harvey and Hodges[25] used $k=1$. The second difference results from the fact that the old figure used a fit to the potential of Diep and Johnson[22] prior to their Erratum; their corrected potential does not disagree as badly with the data as their original one. [We note in passing that the curves labeled "Semi-Classical" in Diep and Johnson's[22] low-temperature figures (Fig. 9 in their original paper and Fig. 2 in the Erratum) appear to be incorrect—the values of the curves seem to correspond to classical calculations.]

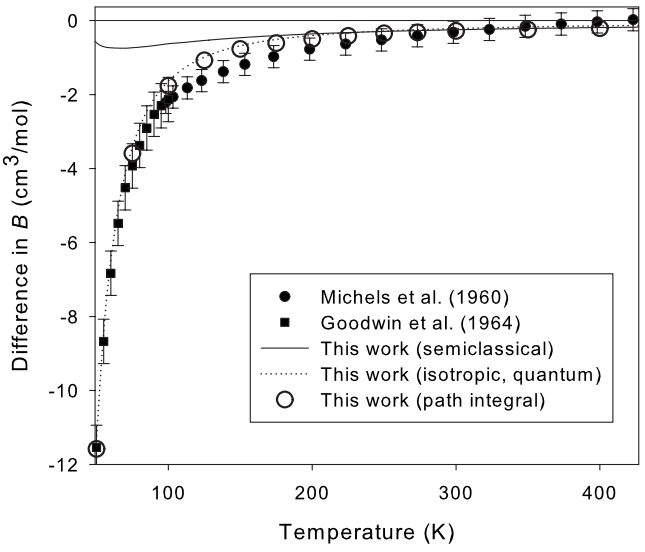

FIG. 12. Comparison of experimental data for the second virial coefficient $B(T)$ with calculated values obtained in this work at various levels of approximation. The baseline for the graph is a semiclassical calculation using the potential of Diep and Johnson (Ref. 22).

We have added to the figure the semiclassical calculation with our new potential, the PIMC calculation with our potential, and the fully quantum computation of $B(T)$ for normal hydrogen from the isotropic approximation to our potential. We have also added points reported by Goodwin *et al.*[70] for *para*-H_2 (also shown with $k=2$ uncertainties) where these data overlap with our calculations.

It can be seen that, when the quantum effects are fully incorporated with the PIMC method, the agreement of $B(T)$ with experimental data is very good. The error observed by Harvey and Hodges[25] came from both of their hypothesized sources. The solid line in Fig. 12 shows the difference between values of $B(T)$ predicted by our potential and by the (revised) DJ potential (both in the semiclassical calculation), and it accounts for most of the deviation in $B(T)$ above about 200 K. However, at lower temperatures, it is clear that the semiclassical calculation is not sufficiently accurate (in fact, as can be seen in Table VI, it diverges unphysically at very low temperatures) and the inclusion of higher-order quantum effects with PIMC is needed to obtain agreement with experiment.

We can also examine the effect of the anisotropy of the potential on $B(T)$. This effect is simply the difference between our PIMC points (open circles in Fig. 12) and the fully quantum calculations with the isotropic potential (dotted line in Fig. 12). It seems that the isotropic potential, when quan-

tum effects are fully accounted for, comes very close to capturing the correct behavior of $B(T)$. Ignoring the anisotropy introduces a small positive bias to $B(T)$.

The source of the small remaining differences with experiment is not clear. At low temperatures, some error must be introduced by our neglect of spin statistics in the PIMC calculation. However, PIMC calculations for *para*-H_2 at 75, 100, and 200 K produced negligible differences from our calculations with no spin restriction. The fully quantum calculations with the isotropic potential were performed for both normal hydrogen and *para*-H_2 for all temperatures, and the differences between the two were similarly small. Even at the lowest temperature of 15 K, the normal-*para* difference was only 0.5 cm^3/mol. At high temperatures, some error is introduced because our potential is for H_2 in its ground state. At higher temperatures, increasingly higher rotational states will be occupied, causing centrifugal stretching. This increase in size of the molecule will cause $B(T)$ to be higher than it would otherwise be, which is at least consistent with the direction of the trend at the high-temperature end of Fig. 12.

VI. CONCLUSIONS

A four-dimensional interaction potential was developed for the H_2-H_2 dimer, as well as its one-dimensional version obtained by averaging over the angles (such a version is needed for quantum calculations of transport properties). This potential is about an order of magnitude more accurate than previously available potentials for this system and is near the limit of what can be achieved with current computational capabilities. A further increase in accuracy of the potential would be costly. The orbital calculations can be performed in basis sets with the cardinal number larger by one, as we did for the test points, as long as sufficient symmetry is present. However, for the configurations with no symmetry, such calculations would be extremely memory and time consuming, especially at the FCI level. Another option is to extend the ECG calculations, but this in turn would require huge amounts of computer and human times to optimize the wave functions. A more realistic possibility for developing a more accurate potential would be to use Gaussian-geminal CCSD and SAPT codes to completely saturate the bulk of the interaction energy, as was done in Refs. 3, 4, 7, 8, 11, and 12 for He_2.

The H_2-H_2 potential was used in semiclassical and PIMC calculations of the second virial coefficient. In contrast to earlier results of Diep and Johnson[22] (obtained with an isotropic version of their potential), we found that the full quantum approach becomes important, giving corrections comparable to experimental uncertainties already for temperatures close to 200 K. Our PIMC results agree with experiment to within the joint uncertainties (at 95% confidence level) for most points in the range 50–600 K. For a couple of points where this is not the case, the computed values almost touch the experimental error bars.

Calculations of transport properties could further validate the potential. Very recently, Mehl[80] performed a fully quantum calculation of the low-density limit of the viscosity from the isotropic form of our potential using the methodology described by Hurly and Mehl.[10] The calculated viscosities are in excellent agreement with the recent precise measurements of May *et al.*[81] which extend from 213 to 394 K. The measured viscosities all fall in the range between 0.01% and 0.06% below the computed values. Viscosities calculated from the isotropic form of the DJ potential deviate from the data by about 0.6% over the same range.

When this work was nearly completed, a paper by Hinde[21] was published presenting a six-dimensional potential energy surface for H_2-H_2. The surface was fitted to *ab initio* CCSD(T) interaction energies computed in the aug-cc-pVQZ+(332) basis, i.e., a basis set one cardinal number smaller than used by us and also with a smaller set of bond functions. No extrapolations were applied in Ref. 21. As Table I shows, the aug-cc-pVQZ+(332) basis set has an uncertainty of 0.39 K at the minimum compared to the CCSD(T) energy used in our fit. Hinde[21] also calculated CCSDT (the coupled-cluster method with the complete account of single, double, and triple excitations) interaction energies on a subset of points using the aug-cc-pVTZ basis and no extrapolations. The CCSDT method gives results close to FCI for H_2-H_2: we have checked that for the minimum, CCSDT recovers about 90% of the contribution beyond CCSD(T). Thus, the use of CCSDT instead of FCI introduces an uncertainty of about 0.2 K in the minimum region. An additional uncertainty at the post-CCSD(T) level arises from the fact that the basis set used by Hinde[21] at this level is smaller than the basis used by us. Although the surface of Hinde[21] is less accurate than that developed by us, it has the important advantage of being six-dimensional. Thus, the two surfaces complement each other very well.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Standards and Technology (Contract Nos. RA1341-07-SE-3530 and RA1341-07-SE-3791) and by NSF Grant No. CHE-0555979. The path-integral calculations of the second virial coefficients were performed on the HPC facility Wiglaf at the Physics Department of the University of Trento. We thank J. K. Johnson and M. R. Moldover for facilitating the collaborations that made this work possible.

[1] T. van Mourik and J. H. van Lenthe, J. Chem. Phys. **102**, 7479 (1995).
[2] R. A. Aziz, A. R. Janzen, and M. R. Moldover, Phys. Rev. Lett. **74**, 1586 (1995).
[3] H. L. Williams, T. Korona, R. Bukowski, B. Jeziorski, and K. Szalewicz, Chem. Phys. Lett. **262**, 431 (1996).
[4] T. Korona, H. L. Williams, R. Bukowski, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **106**, 5109 (1997).
[5] A. R. Janzen and R. A. Aziz, J. Chem. Phys. **107**, 914 (1997).
[6] J. J. Hurly and M. R. Moldover, J. Res. Natl. Inst. Stand. Technol. **105**, 667 (2000).
[7] M. Jeziorska, R. Bukowski, W. Cencek, M. Jaszuński, B. Jeziorski, and K. Szalewicz, Collect. Czech. Chem. Commun. **68**, 463 (2003).
[8] W. Cencek, M. Jeziorska, R. Bukowski, M. Jaszuński, B. Jeziorski, and K. Szalewicz, J. Phys. Chem. A **108**, 3211 (2004).
[9] J. B. Anderson, J. Chem. Phys. **120**, 9886 (2004).
[10] J. J. Hurly and J. B. Mehl, J. Res. Natl. Inst. Stand. Technol. **112**, 75 (2007).
[11] K. Patkowski, W. Cencek, M. Jeziorska, B. Jeziorski, and K. Szalewicz, J. Phys. Chem. A **111**, 7611 (2007).
[12] M. Jeziorska, W. Cencek, K. Patkowski, B. Jeziorski, and K. Szalewicz,

J. Chem. Phys. **127**, 124303 (2007).
[13] W. Cencek, M. Jeziorska, O. Akin-Ojo, and K. Szalewicz, J. Phys. Chem. A **111**, 11311 (2007).
[14] R. Hellmann, E. Bich, and E. Vogel, Mol. Phys. **105**, 3013 (2007).
[15] E. Bich, R. Hellmann, and E. Vogel, Mol. Phys. **105**, 3035 (2007).
[16] S. Grebenev, B. Sartakov, J. Toennies, and A. Vilesov, Science **289**, 1532 (2000).
[17] S. Moroni, M. Botti, S. De Palo, and A. R. W. McKellar, J. Chem. Phys. **122**, 094314 (2005).
[18] Y. Kwon and K. B. Whaley, Phys. Rev. Lett. **89**, 273401 (2002).
[19] F. Paesani and K. B. Whaley, J. Chem. Phys. **124**, 234310 (2006).
[20] K. Szalewicz, Int. Rev. Phys. Chem. **27**, 273 (2008).
[21] R. J. Hinde, J. Chem. Phys. **128**, 154308 (2008).
[22] P. Diep and J. K. Johnson, J. Chem. Phys. **112**, 4465 (2000); **113**, 3480(E) (2000).
[23] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[24] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[25] A. H. Harvey and M. P. Hodges, J. Chem. Phys. **113**, 3474 (2000).
[26] A. I. Boothroyd, P. Martin, W. J. Keogh, and M. J. Peterson, J. Chem. Phys. **116**, 666 (2002).
[27] T.-G. Lee, N. Balakrishnan, R. C. Forrey, P. Stancil, D. R. Schultz, and G. J. Ferland, J. Chem. Phys. **125**, 114302 (2006).
[28] S. Bock, E. Bich, E. Vogel, A. S. Dickinson, and V. Vesovic, J. Chem. Phys. **117**, 2151 (2002).
[29] S. Bock, E. Bich, E. Vogel, A. S. Dickinson, and V. Vesovic, J. Chem. Phys. **120**, 7987 (2004).
[30] M. Takahashi and M. Imada, J. Phys. Soc. Jpn. **53**, 3765 (1984).
[31] G. K. Schenter, J. Chem. Phys. **117**, 6573 (2002).
[32] W. Kolos, K. Szalewicz, and H. Monkhorst, J. Chem. Phys. **84**, 3278 (1986).
[33] R. J. Hinde, Few-Body Syst. **38**, 187 (2006).
[34] R. J. LeRoy and J. M. Hutson, J. Chem. Phys. **86**, 837 (1987).
[35] P. Jankowski and K. Szalewicz, J. Chem. Phys. **123**, 104301 (2005).
[36] R. Lindh, F. R. Manby, M. Schütz *et al.*, MOLPRO, a package of *ab initio* programs designed by H.-J. Werner and P. J. Knowles, version 2006.1 (see http://www.molpro.net).
[37] Certain commercial products are identified in this paper but only in order to adequately specify the procedure. Such identification neither constitutes nor implies recommendation or endorsement by either the U.S. government or the National Institute of Standards and Technology.
[38] J. Olsen with contributions from H. Larsen and M. Fulscher, LUCIA, A full CI, general active space program.
[39] R. Bukowski, W. Cencek, P. Jankowski, M. Jeziorska, B. Jeziorski, S. A. Kucharski, V. F. Lotrich, A. J. Misquitta, R. Moszyński, K. Patkowski, R. Podeszwa, S. Rybak, K. Szalewicz, H. L. Williams, R. J. Wheatley, P. E. S. Wormer, and P. S. Żuchowski, SAPT2006, an *ab initio* program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies, University of Delaware and University of Warsaw (http://www.physics.udel.edu/~szalewic/SAPT/SAPT.html).
[40] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
[41] S. L. Mielke, B. C. Garrett, and K. A. Peterson, J. Chem. Phys. **111**, 3806 (1999).
[42] S. L. Mielke, B. C. Garrett, and K. A. Peterson, J. Chem. Phys. **116**, 4142 (2002).
[43] S. L. Mielke and K. A. Peterson (unpublished).
[44] R. N. Hill, J. Chem. Phys. **83**, 1173 (1985).
[45] W. Kutzelnigg, Theor. Chim. Acta **68**, 445 (1985).
[46] W. Kutzelnigg and J. D. Morgan III, J. Chem. Phys. **96**, 4484 (1992).
[47] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).
[48] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).
[49] M. Jeziorska, W. Cencek, K. Patkowski, B. Jeziorski, and K. Szalewicz, Int. J. Quantum Chem. **108**, 2053 (2008).
[50] D. S. Feller, J. Chem. Phys. **96**, 6104 (1992).
[51] F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. H. van Lenthe, Chem. Rev. (Washington, D.C.) **94**, 1873 (1994).
[52] R. Bukowski, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **104**, 3306 (1996).
[53] K. Szalewicz and B. Jeziorski, J. Chem. Phys. **109**, 1198 (1998).
[54] See EPAPS Document No E-JCPSA6-129-001834 for the list of computed interaction energies, fit parameters, and a FORTRAN subroutine calculating the fit. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[55] DALTON, a molecular electronic structure program, Release 2.0, 2005 (see http://www.kjemi.uio.no/software/dalton/dalton.html).
[56] B. Jeziorski, R. Moszyński, and K. Szalewicz, Chem. Rev. (Washington, D.C.) **94**, 1887 (1994).
[57] K. Szalewicz, K. Patkowski, and B. Jeziorski, Struct. Bonding (Berlin) **116**, 43 (2005).
[58] K. Patkowski, K. Szalewicz, and B. Jeziorski, J. Chem. Phys. **125**, 154107 (2006).
[59] K. Patkowski and K. Szalewicz, J. Chem. Phys. **127**, 164103 (2007).
[60] J. Rychlewski and J. Komasa, in *Explicitly Correlated Wave Functions in Chemistry and Physics: Theory and Applications*, edited by J. Rychlewski (Kluwer, Dordrecht, 2003), p. 91.
[61] W. Cencek, J. Komasa, K. Pachucki, and K. Szalewicz, Phys. Rev. Lett. **95**, 233004 (2005).
[62] W. Cencek and K. Szalewicz, Int. J. Quantum Chem. **108**, 2191 (2008).
[63] W. Cencek, J. Komasa, and J. Rychlewski, Chem. Phys. Lett. **246**, 417 (1995).
[64] R. Bukowski, J. Sadlej, B. Jeziorski, P. Jankowski, K. Szalewicz, S. A. Kucharski, H. L. Williams, and B. M. Rice, J. Chem. Phys. **110**, 3785 (1999).
[65] K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).
[66] D. Spelsberg, T. Lorenz, and W. Meyer, J. Chem. Phys. **99**, 7845 (1993).
[67] D. M. Bishop and J. Pipin, Int. J. Quantum Chem. **45**, 349 (1993).
[68] J. Komasa and A. J. Thakkar, Mol. Phys. **78**, 1039 (1993).
[69] M. P. Hodges, R. J. Wheatley, G. K. Schenter, and A. H. Harvey, J. Chem. Phys. **120**, 710 (2004).
[70] R. D. Goodwin, D. E. Diller, H. M. Roder, and L. A. Weber, J. Res. Natl. Bur. Stand. **68A**, 121 (1964).
[71] R. D. Goodwin, D. E. Diller, H. M. Roder, and L. A. Weber, J. Res. Natl. Bur. Stand. **67A**, 173 (1963).
[72] A. Michels, W. de Graaff, and C. A. ten Seldam, Physica (Amsterdam) **26**, 393 (1960).
[73] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954).
[74] G. Garberoglio, J. Chem. Phys. **128**, 134109 (2008).
[75] D. Marx and M. H. Müser, J. Phys.: Condens. Matter **11**, R117 (1999).
[76] L. D. Fosdick and H. F. Jordan, Phys. Rev. **143**, 58 (1966).
[77] S. Miura, J. Chem. Phys. **126**, 114308 (2007).
[78] G. Garberoglio and J. K. Johnson, "Hybrid path integral Monte Carlo simulation of rigid diatomic molecules: Effect of quantized rotations on the selectivity of hydrogen isotopes in carbon nanotubes," e-print arXiv:cond-mat/0703287.
[79] R. T Pack, J. Chem. Phys. **78**, 7217 (1983).
[80] J. B. Mehl (unpublished).
[81] E. F. May, R. F. Berg, and M. R. Moldover, Int. J. Thermophys. **28**, 1085 (2007).