The Journal of Chemical Physics

Highly accurate potential energy surface for the He–H_2 dimer

Cite as: J. Chem. Phys. **139**, 144305 (2013); https://doi.org/10.1063/1.4824299
Submitted: 05 July 2013 . Accepted: 20 September 2013 . Published Online: 08 October 2013

Brandon W. Bakr, Daniel G. A. Smith, and Konrad Patkowski

Highly accurate potential energy surface for the He–H_2 dimer

Brandon W. Bakr, Daniel G. A. Smith, and Konrad Patkowski[a)]
Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA

(Received 5 July 2013; accepted 20 September 2013; published online 8 October 2013)

A new highly accurate interaction potential is constructed for the He–H_2 van der Waals complex. This potential is fitted to 1900 *ab initio* energies computed at the very large-basis coupled-cluster level and augmented by corrections for higher-order excitations (up to full configuration interaction level) and the diagonal Born-Oppenheimer correction. At the vibrationally averaged H–H bond length of 1.448736 bohrs, the well depth of our potential, 15.870 ± 0.065 K, is nearly 1 K larger than the most accurate previous studies have indicated. In addition to constructing our own three-dimensional potential in the van der Waals region, we present a reparameterization of the Boothroyd-Martin-Peterson potential surface [A. I. Boothroyd, P. G. Martin, and M. R. Peterson, J. Chem. Phys. **119**, 3187 (2003)] that is suitable for all configurations of the triatomic system. Finally, we use the newly developed potentials to compute the properties of the lone bound states of ^{4}He–H_2 and ^{3}He–H_2 and the interaction second virial coefficient of the hydrogen-helium mixture. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4824299]

I. INTRODUCTION

Hydrogen and helium are the two most abundant species in the universe. Accordingly, the interactions between helium and molecular hydrogen are critical for understanding the properties of many diverse astrophysical environments, from the cooling of primordial gas to form stars[1,2] all the way to the spectral features of ultracool white dwarfs.[3] In the laboratory, the He–H_2 interactions are a significant factor determining the properties of hydrogen clusters inside helium droplets (a possible way to supercool liquid *para*-H_2 below its triple point of 13.8 K down to its superfluid transition temperature of ∼1 K[4,5]) and the formation of ultracold H_2 molecules by cooling with helium atoms in an optical trap.[6]

The other factor besides the experimental importance that has contributed to the large number of computational studies for the He–H_2 complex is its simplicity. With only four electrons, He–H_2 is the simplest atom-molecule dimer. Thus, it is a suitable testing ground for assessing the accuracy of different approaches to obtain *ab initio* weak interaction energies, fit analytical potentials to these energies, and compute properties of bound and resonance states as well as elastic and inelastic scattering cross sections. Among many *ab initio* He–H_2 potentials constructed to date,[7–14] the Boothroyd-Martin-Peterson (BMP) surface[14] is generally considered the most accurate. This elaborate potential, containing 112 adjustable parameters, was fitted to 25 065 *ab initio* data points representing all kinds of conformations of the triatomic system, including ones with very short (repulsive) or very long (dissociative) values of the hydrogen-hydrogen distance r_{HH}. Thus, an important virtue of the BMP potential is its ability to describe all phenomena associated with the He–H–H triatomic system, including high-energy and reactive scattering. On the other hand, the van der Waals (vdW) well region of the BMP potential energy surface (PES) was described by data points computed in earlier *ab initio* studies.[11–13] The most accurate of them, the 69 points of Tao,[13] were obtained using the fourth-order Møller-Plesset perturbation theory (MP4) and a 6*s*4*p*2*d* + (3*s*3*p*2*d*) basis set (we will adopt a notation where the midbond functions, centered halfway between the H_2 center of mass (COM) and the helium atom, are listed in parentheses).

The BMP potential has been employed in several studies of the bound and resonance rovibrational states of the He–H_2 complex[15–18] and of the low-energy elastic and inelastic collisions of different isotopomers of He–H_2.[17,19] In the latter case, Ref. 19 found that the computed state-to-state rovibrational cross sections were in quite a poor agreement with experimental data:[20] in fact, an older Muchnick-Russek (MR) potential[12] provided better agreement and therefore was used in subsequent scattering studies.[21,22]

It should be noted that the methods and basis sets employed in the *ab initio* calculations utilized to fit the BMP potential are far from the current state of the art. Since 1994, when the MR[12] and Tao[13] calculations were published, the coupled-cluster method with single, double, and perturbative triple excitations [CCSD(T)][23] has been unambiguously established as the "gold standard" of quantum chemistry. The use of CCSD(T), combined with correlation-consistent basis sets,[24] complete-basis-set (CBS) extrapolations,[25] midbond functions[26] and, more recently, explicitly correlated R12/F12 approaches,[27,28] has allowed for calculations of weak interaction energies of small and medium closed-shell dimers with high and consistent accuracy.[29,30] In the last few years, the importance of effects beyond CCSD(T), in particular of quadruple excitations at the perturbative CCSDT(Q) level,[31,32] has been investigated for a number of atomic and small molecular dimers.[33–38] For four-electron systems like He–H_2, CCSDTQ is equivalent to full configuration interaction (FCI), and an

a)Author to whom correspondence should be addressed. Electronic mail: patkowsk@auburn.edu

exact inclusion of quadruple excitations in a moderate basis set is feasible.

A comparison of different four-electron complexes indicates that computing an *ab initio* He–H_2 potential is more demanding than for He–He but easier than for H_2–H_2. There are three factors contributing to this: the number of basis function centers, the point group symmetry, and the number of data points required to fit an analytical potential of a given dimensionality. The one-dimensional (1D) potential for the helium dimer is now known to a truly amazing accuracy: at the minimum separation, the nonrelativistic Born-Oppenheimer (BO) He–He interaction energy amounts to -11.0006 ± 0.0002 K[39] as calculated variationally using four-electron explicitly correlated functions.[40] The relativistic, adiabatic, and quantum electrodynamics corrections to the He–He potential have been computed to similar accuracy.[41] The H_2–H_2 potential is 4D for rigid monomers and 6D for flexible monomers. In the former case, 213 *ab initio* data points were computed[42] using CCSD(T) in basis sets as large as aug-cc-pV5Z+(3*s*3*p*2*d*2*f*1*g*)≡a5Z+(33221) plus a FCI correction in bases as large as aTZ+(332), with both contributions extrapolated to the CBS limit. The minimum H_2–H_2 interaction energy obtained in this way amounted to -56.96 ± 0.16 K. The available 6D potentials are only slightly less accurate.[43,44] These results suggest that obtaining a He–H_2 potential accurate to a few hundredths of a Kelvin at the minimum should be feasible using state-of-the-art electronic structure theory. As we will see below, the potential of Tao,[13] and thus also the BMP potential, is only accurate to about 1 K at the minimum, an error large enough to account for the observed discrepancies of the BMP and experimental scattering cross sections.[19] Interestingly, an accurate, CCSD(T)/a5Z He–H_2 dipole moment surface has been constructed recently,[45] but no effort has been made to improve the BMP PES.

The objective of the present work is to construct a highly accurate He–H_2 interaction potential using state-of-the-art techniques. The bulk of the interaction energy will be recovered at the CCSD(T) level, but the contributions from higher coupled-cluster excitations (all the way through FCI), relativistic corrections, and effects beyond the BO approximation will also be investigated. Large augmented correlation-consistent a*XZ* bases with midbond functions, coupled with CBS extrapolations, will be employed to ensure that the results are virtually converged with the basis set. The resulting *ab initio* data will be fitted to an analytical 3D potential providing a greatly improved description of the He–H_2 vdW well compared to BMP. As our primary goal is reproducing bound-state and low-energy scattering properties accurately, we will make no effort to improve the regions of the surface where r_{HH} is far from its equilibrium value of 1.4015 bohrs.[46] However, in addition to generating our own 3D fit for the van der Waals well, we will refit the 112-parameter BMP potential replacing the original *ab initio* data in the vdW region by our, significantly more accurate, values. The resulting potential combines an improved account of the vdW well with a BMP-level description of other surface regions. Our new potential will then be employed to compute the single bound vibrational state of ^{4}He–H_2 and ^{3}He–H_2 and the second virial coefficient of the hydrogen-helium mixture.

The structure of the rest of this paper is as follows. The methodology and computational details concerning the calculations of *ab initio* data points are described in Sec. II. The resulting interaction energies are presented in Sec. III. Section IV describes our analytical fit to the *ab initio* data and Sec. V presents our reparameterization of the BMP surface. Section VI details the properties of the bound rovibrational state. Section VII describes the calculation of the second virial coefficient. Finally, Sec. VIII presents conclusions.

II. METHODS AND COMPUTATIONAL DETAILS

The geometry of the interacting He–H_2 system can be specified by three variables: the H–H bond length r_{HH}, the distance R from the helium atom to the center of the H_2 molecule, and the angle θ between the H_2 molecular axis and the line connecting the helium atom with the H_2 COM. Due to the symmetry of the system, only the values $0^\circ \le \theta \le 90^\circ$ need to be considered. The *ab initio* calculations have been carried out on a 1900-point (r_{HH}, R, θ) grid, where $r_{HH} \in \{1.1, 1.2, 1.3, 1.35, 1.4, 1.448736, 1.5, 1.55, 1.65, 1.75\}$ bohrs, $R \in \{3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.2, 6.3, 6.4, 6.5, 6.6, 6.8, 7.0, 7.5, 8.0, 9.0, 10.0, 12.0, 15.0\}$ bohrs, and $\theta \in \{0^\circ, 10^\circ, 20^\circ, 30^\circ, 40^\circ, 50^\circ, 60^\circ, 70^\circ, 80^\circ, 90^\circ\}$. Thus, the calculations cover the entire van der Waals well and the repulsive region up to about 3000 K, and the H_2 molecule can undergo vibrations but cannot break apart. The value $\langle r_{HH} \rangle = 1.448736$ bohrs represents the H–H distance averaged over the ground vibrational state of H_2.[47,48]

The interaction energy E_{int} for a given geometry (r_{HH}, R, θ) is a sum of the monomer deformation energy $E_{mon,def} = E_{H_2}(r_{HH}) - E_{H_2}(r_0)$ (where $r_0 = 1.4015$ bohrs is the equilibrium bond length of H_2[46]) and the two-body interaction energy

$$E_{int,2B} = E_{int}^{CCSD(T)} + \Delta E_{int}^{CCSDT} + \Delta E_{int}^{FCI} + \Delta E_{int}^{DBOC}, \quad (1)$$

where all "int" quantities on the rhs are interaction energy contributions ($E_{int}^X = E_{He-H_2}^X - E_{He}^X - E_{H_2}^X$ at any level X), $\Delta E_{int}^{CCSDT} = E_{int}^{CCSDT} - E_{int}^{CCSD(T)}$, $\Delta E_{int}^{FCI} = E_{int}^{FCI} - E_{int}^{CCSDT}$, and ΔE_{int}^{DBOC} is an estimate of the diagonal Born-Oppenheimer correction (DBOC) term. The relativistic correction to interaction energy, ΔE_{int}^{rel}, was also considered but turned out to be negligible.

The calculations employed standard singly and doubly augmented correlation-consistent basis sets aug-cc-pV*XZ*≡a*XZ* and d-aug-cc-pV*XZ*≡da*XZ* with X = T, Q, 5, 6, and 7. The basis sets through $X = 6$ have been optimized in Refs. 24 (hydrogen) and 49 (helium). The $X = 7$ bases were taken from Ref. 50 (the aug-mcc-pV7Z set) for hydrogen and Ref. 51 for helium. Two kinds of midbond sets were placed halfway between the He atom and the COM of the H_2 molecule: a "constant midbond" set (3*s*3*p*2*d*2*f*) ≡ (3322), employed before in several calculations of weak interaction energies[52,53] (the exponents are 0.9, 0.3, and 0.1 for *sp* and 0.6, 0.2 for *df* functions), and a "variable midbond" set a*XZ*, taken as the hydrogen-atom basis with X the same as for the atom-centered part of the set. The inclusion of the (a*XZ*) midbond increases the basis set size by one third while the

TABLE I. Two-body interaction energy contributions (in Kelvin) for the near-minimum geometry (r_{HH}, R, θ) = (1.448736 bohrs, 6.4 bohrs, 0°) computed using different basis sets. The columns marked "extr." contain CBS-extrapolated contributions, with the value in the X row obtained using basis sets with cardinal numbers $X-1$ and X. The values at the level used to compute all 1900 *ab initio* data points are marked in bold. The CCSD(T) calculations for the da5Z+(da5Z) and da7Z+(da7Z) bases failed to converge.

	$E_{int}^{CCSD(T)}$		ΔE_{int}^{CCSDT}		ΔE_{int}^{FCI}	
Basis set	Computed	extr.	Computed	extr.	Computed	extr.
aTZ	−13.234		−0.581		**−0.025**	
aQZ	−14.251	−15.001	−0.568	−0.558	−0.028	−0.031
a5Z	−14.742	−15.199	−0.557	**−0.545**	−0.029	−0.030
a6Z	−15.001	−15.327	−0.551	−0.544		
a7Z	−15.104	−15.285				
aTZ+(3322)	−14.923		−0.581		−0.029	
aQZ+(3322)	−15.106	−15.218	−0.566	−0.556	−0.029	−0.029
a5Z+(3322)	−15.155	−15.206	−0.555	−0.544		
a6Z+(3322)	−15.184	−15.221	−0.550	−0.542		
a7Z+(3322)	−15.199	−15.226				
aTZ+(aTZ)	−14.801		−0.586		−0.028	
aQZ+(aQZ)	−15.103	−15.327	−0.566	−0.551	−0.029	−0.030
a5Z+(a5Z)	−15.175	−15.252	−0.554	−0.541		
a6Z+(a6Z)	−15.206	**−15.249**				
a7Z+(a7Z)	−15.218	−15.239				
daTZ	−13.735					
daQZ	−14.738	−15.465				
da5Z	−15.022	−15.310				
da6Z	−15.141	−15.301				
da7Z	−15.181	−15.253				
daTZ+(3322)	−14.953					
daQZ+(3322)	−15.138	−15.252				
da5Z+(3322)	−15.173	−15.215				
da6Z+(3322)	−15.196	−15.225				
da7Z+(3322)	−15.209	−15.230				
daTZ+(daTZ)	−14.869					
daQZ+(daQZ)	−15.138	−15.344				
da6Z+(da6Z)	−15.213					

addition of the (3322) set enlarges the basis by 6% (a7Z)–52% (aTZ). The addition of midbond functions has been shown to significantly speed up basis set convergence of dispersion-dominated interaction energies[26,54] and can be efficiently combined with standard CBS extrapolations[55] even though the constant-midbond and variable-midbond sequences exhibit slightly different convergence patterns.[56] All interaction energies presented in this work include the full counterpoise (CP) correction for basis set superposition error.

The relativistic corrections were estimated at the CCSD(T) level of theory using the second-order Douglas-Kroll-Hess (DKH) Hamiltonian[57,58] and the a6Z+(a6Z) basis. The DBOC correction was computed at the CCSD level[59] using standard a*X*Z basis sets with and without midbond.

All CCSD(T) energies and relativistic corrections were computed with the MOLPRO2010.1 code.[60] The CCSDT energies as well as the DBOC terms were obtained using the CFOUR code,[61] and the FCI energies were calculated by the LUCIA program[62] interfaced to DALTON2.0.[63] The He and H_2 monomers have only two electrons each and CCSD is equivalent to FCI. Therefore, the CCSDT and FCI corrections were computed for the dimer only.

III. INTERACTION ENERGIES

Our first task is to determine which basis sets give the highest accuracy at a given level of theory while being feasible for all data points. To this end, we selected two characteristic points: (r_{HH}, R, θ) = (1.448736 bohrs, 6.4 bohrs, 0°) (near the vdW minimum) and (r_{HH}, R, θ) = (1.448736 bohrs, 6.3 bohrs, 90°) (near the saddle point in the minimum-energy valley around the H_2 molecule). For these two points, we ran an extensive set of CCSD(T), CCSDT, and FCI calculations in different basis sets with and without midbond. The results are collected in Tables I and II for the minimum and saddle point geometries, respectively. In addition to the computed results, we present the values extrapolated to the CBS limit using the standard X^{-3} technique.[25] Only the correlation part of the interaction energy is extrapolated; the Hartree-Fock (HF) part converges significantly faster with the basis set and this part was taken straight from the calculation in the larger of the two basis sets used in the extrapolation. For example, the energy

TABLE II. Two-body interaction energy contributions (in Kelvin) for the near-saddle point geometry (r_{HH}, R, θ) = (1.448736 bohrs, 6.3 bohrs, 90°) computed using different basis sets. The columns marked "extr." contain CBS-extrapolated contributions, with the value in the X row obtained using basis sets with cardinal numbers $X-1$ and X. The values at the level used to compute all 1900 *ab initio* data points are marked in bold.

	$E_{int}^{CCSD(T)}$		ΔE_{int}^{CCSDT}		ΔE_{int}^{FCI}	
Basis set	Computed	extr.	Computed	extr.	Computed	extr.
aTZ	−11.022		−0.446		**−0.019**	
aQZ	−12.369	−13.304	−0.441	−0.437	−0.022	−0.024
a5Z	−13.010	−13.667	−0.434	**−0.427**	−0.023	−0.024
a6Z	−13.357	−13.846	−0.431	−0.427		
a7Z	−13.550	−13.878				
aTZ+(3322)	−13.463		−0.454		−0.021	
aQZ+(3322)	−13.608	−13.677	−0.444	−0.436	−0.023	−0.024
a5Z+(3322)	−13.657	−13.700	−0.435	−0.425		
a6Z+(3322)	−13.686	−13.726	−0.430	−0.423		
a7Z+(3322)	−13.703	−13.731				
aTZ+(aTZ)	−13.291		−0.457		−0.021	
aQZ+(aQZ)	−13.605	−13.813	−0.444	−0.434	−0.023	−0.024
a5Z+(a5Z)	−13.676	−13.741	−0.434	−0.423		
a6Z+(a6Z)	−13.708	**−13.750**				
a7Z+(a7Z)	−13.721	−13.743				
daTZ	−11.738					
daQZ	−13.124	−14.116				
da5Z	−13.495	−13.886				
da6Z	−13.634	−13.825				
da7Z	−13.681	−13.759				
daTZ+(3322)	−13.508					
daQZ+(3322)	−13.641	−13.704				
da5Z+(3322)	−13.679	−13.713				
da6Z+(3322)	−13.700	−13.727				
da7Z+(3322)	−13.713	−13.734				
daTZ+(daTZ)	−13.371					
daQZ+(daQZ)	−13.645	−13.825				
da5Z+(da5Z)	−13.698	−13.744				
da6Z+(da6Z)	−13.717	−13.742				
da7Z+(da7Z)	−13.725	−13.740				

$E_{int}^{CCSD(T)}$ extrapolated from the a5Z and a6Z bases, which will be denoted by $E_{int}^{CCSD(T)}$/a(5,6)Z, is a sum of the HF part computed in the a6Z basis and the correlation part obtained from the a5Z and a6Z results via the X^{-3} extrapolation.

The simplicity of the He–H_2 dimer allows for CCSD(T) calculations in the largest available basis sets, all the way to doubly augmented septuple zeta (although, for the near-minimum geometry, the da5Z+(da5Z) and da7Z+(da7Z) calculations failed to converge due to near linear dependencies in the basis set). The largest sets we were able to use at the CCSDT and FCI levels were a6Z+(3322) (417 functions) and a5Z (240 functions), respectively. The results in Tables I and II show that CCSD(T) recovers 96%–97% of the FCI interaction energy and that the inclusion of full triples increases this percentage to 99.8%. Moreover, it is evident that the presence of midbond functions dramatically improves basis set convergence of the CCSD(T) interaction energy. Without midbond, even the a7Z result is not fully converged and a similar accuracy can be attained with the addition of a midbond (either constant or variable) to the aQZ basis set. While the addition of a second augmentation obviously improves the CCSD(T) interaction energies, it is not cost effective because a larger improvement is provided by the midbond functions. We chose to use conventional CCSD(T) in favor of the explicitly correlated approaches CCSD(T)-F12a, CCSD(T)-F12b,[64,65] or CCSD(T)(F12*)[66] because, at this level of basis set convergence, the residual approximations to CCSD-F12 and a conventional (non-F12) treatment of triples become limiting factors for the accuracy attainable with the F12 approach.[56,67] Obviously, still more accurate interaction energies could be obtained using four-electron explicitly correlated functions with fully optimized nonlinear parameters.[68] Such calculations have been carried out for selected geometries of the H_2–H_2 complex,[42,69] however, they are too expensive to perform on the entire He–H_2 PES.

The ΔE_{int}^{CCSDT} correction in Tables I and II exhibits fast basis set convergence and is already accurate to about 20 mK at the aQZ level. Contrary to the $E_{int}^{CCSD(T)}$ case, the addition of midbond functions is not beneficial for the full triples correction. The ΔE_{int}^{FCI} effect is very small and converges even faster in absolute terms—the aTZ results are already converged to a few mK.

TABLE III. The diagonal Born-Oppenheimer correction to the He–H_2 interaction energy (in Kelvin) for the near-minimum and near-saddle point geometries computed using different basis sets.

Basis set	Minimum	Saddle point
aTZ	−0.051	−0.026
aQZ	−0.054	−0.027
a5Z	−0.055	−0.028
a6Z	−0.055	−0.029
aTZ+(aTZ)	−0.051	−0.028
aQZ+(aQZ)	−0.054	−0.029
a5Z+(a5Z)	−0.054	−0.029

At the level of accuracy pursued in this work, one has to go beyond the nonrelativistic BO treatment. Therefore, we investigated the role of the relativistic correction (estimated using CCSD(T) with the second-order DKH Hamiltonian) and of the DBOC term (computed at the CCSD level) for the near-minimum and near-saddle point interaction energies presented in Tables I and II. The relativistic correction, computed in the a6Z+(a6Z) basis set, amounts to 0.002 K for each of the two geometries. Thus, this correction is negligible and we will exclude it from further consideration. However, as the results in Table III indicate, the DBOC correction is larger and needs to be taken into account for all *ab initio* data points. Fortunately, this correction exhibits fast basis set convergence and computing it in the aTZ basis is fully adequate.

Combining the most accurate estimates of different contributions: $E_{\rm int}^{\rm CCSD(T)}$/a(6,7)Z+(a(6,7)Z), $\Delta E_{\rm int}^{\rm CCSDT}$/a(5,6)Z+(3322), $\Delta E_{\rm int}^{\rm FCI}$/a(Q,5)Z, $\Delta E_{\rm int}^{\rm rel}$/a6Z+(a6Z), and $\Delta E_{\rm int}^{\rm DBOC}$/a6Z, gives total two-body He–H_2 interaction energies of −15.864 ± 0.035 K and −14.217 ± 0.035 K for the near-minimum and near-saddle point geometries, respectively. The uncertainties of the CBS-extrapolated contributions $E_{\rm int}^{\rm CCSD(T)}$, $\Delta E_{\rm int}^{\rm CCSDT}$, and $\Delta E_{\rm int}^{\rm FCI}$ were estimated as differences between the extrapolated result and the value computed in the larger of the two bases. Then, the uncertainties of these contributions were added linearly along with an assumed 0.005 K estimate of the uncertainty for the relativistic and DBOC terms.

Some of the largest-basis calculations in Tables I and II would be too costly to perform for all data points including ones with only C_s point-group symmetry. Therefore, we decided to restrict the calculations for all 1900 geometries to bases up to a6Z+(a6Z), a5Z, aTZ, and aTZ for the $E_{\rm int}^{\rm CCSD(T)}$, $\Delta E_{\rm int}^{\rm CCSDT}$, $\Delta E_{\rm int}^{\rm FCI}$, and $\Delta E_{\rm int}^{\rm DBOC}$ contributions, respectively. Specifically, the *ab initio* two-body interaction energies were estimated as

$$E_{\rm int,2B} = E_{\rm int}^{\rm CCSD(T)}/{\rm a(5,6)Z}+({\rm a(5,6)Z})+\Delta E_{\rm int}^{\rm CCSDT}/{\rm a(Q,5)Z} + \Delta E_{\rm int}^{\rm FCI}/{\rm aTZ} + \Delta E_{\rm int}^{\rm DBOC}/{\rm aTZ}. \quad (2)$$

For the two characteristic points presented in Tables I and II, Eq. (2) leads to two-body interaction energy estimates of −15.870 ± 0.065 K and −14.222 ± 0.059 K, respectively (where we have estimated the uncertainties of $E_{\rm int}^{\rm CCSD(T)}$ and $\Delta E_{\rm int}^{\rm CCSDT}$ as before but using smaller basis sets, and assumed a 0.005 K uncertainty for $\Delta E_{\rm int}^{\rm FCI}$ in addition to the 0.005 K uncertainty for the combined relativistic and DBOC terms). One should note the remarkable consistency of these estimates with the largest-basis-set values of −15.864 ± 0.035 K and −14.217 ± 0.035 K, respectively, indicating that the interaction energies from Eq. (2) are highly accurate and that our uncertainty estimates are reasonable (in fact, quite conservative).

A two-dimensional (R, θ) cross section through the two-body PES corresponding to $r_{\rm HH}$ = 1.448736 bohrs is presented in Fig. 1. This figure shows that the potential is only weakly anisotropic throughout the van der Waals well. There is, however, a somewhat stronger θ dependence in the repulsive region. As $E_{\rm int,\,2B}(r_{\rm HH}, R, \theta) = E_{\rm int,\,2B}(r_{\rm HH}, R, 180^\circ - \theta)$, only the symmetry unique part of the PES is shown in Fig. 1.

The H–H potential energy curve, needed to compute the monomer deformation correction $E_{\rm mon,\,def}$ to the two-body interaction energy, is known to a truly amazing accuracy. In this work, we chose to follow the same approach as in the BMP potential[14] and computed the H–H binding energies from the analytical fit of Schwenke.[46] This fit, constructed in 1988, compares very well to the newer, ultra-accurate *ab initio* results such as the calculations of Wolniewicz[70] that include the relativistic and nonadiabatic corrections. In particular, at the near-minimum separation of 1.4 bohrs, the Schwenke potential gives a binding energy of −0.17449537 hartree, while the *ab initio* value of Ref. 70 is −0.17449581 hartree, a difference of just 0.14 K. The accuracy of the nonrelativistic BO potential and of the relativistic and quantum electrodynamics corrections has been further improved in the last few years,[40,71–74] but the differences with respect to the results of Ref. 70 are very small. For example, the most accurate nonrelativistic BO result at 1.4 bohrs[74] differs from the Ref. 70 value by only 0.014 K. The additional benefit of using the Schwenke potential to account for monomer deformation energy is that any comparison between our potential and the BMP one[14] can be performed on equal terms using total interaction energies $E_{\rm int}$ and their two-body components $E_{\rm int,\,2B}$, cf. Sec. V.

IV. ANALYTICAL FITS

To provide a highly accurate representation of the He–H_2 van der Waals well, a 3D analytical potential was fitted to the 1900 *ab initio* two-body interaction energies. This potential has a form

$$V(r_{\rm HH}, R, \theta) = \sum_{k=0}^{3} r_{\rm HH}^k \sum_{l=0,2,4} (A_{kl0} + A_{kl1} R + A_{kl2} R^2) \times {\rm e}^{-\alpha_k R} P_l(\cos\theta) - \sum_{k=0}^{2} r_{\rm HH}^k \sum_{l=0,2} \sum_{n=6,8,10} f_n(dR) \frac{C_{nkl}}{R^n} P_l(\cos\theta), \quad (3)$$

where P_l is the standard Legendre polynomial (with only even values of l needed because of symmetry) and $f_n(dR)$ is the nth Tang-Toennies damping function[75]

$$f_n(dR) = 1 - {\rm e}^{-dR} \sum_{m=0}^{n} \frac{(dR)^m}{m!}. \quad (4)$$

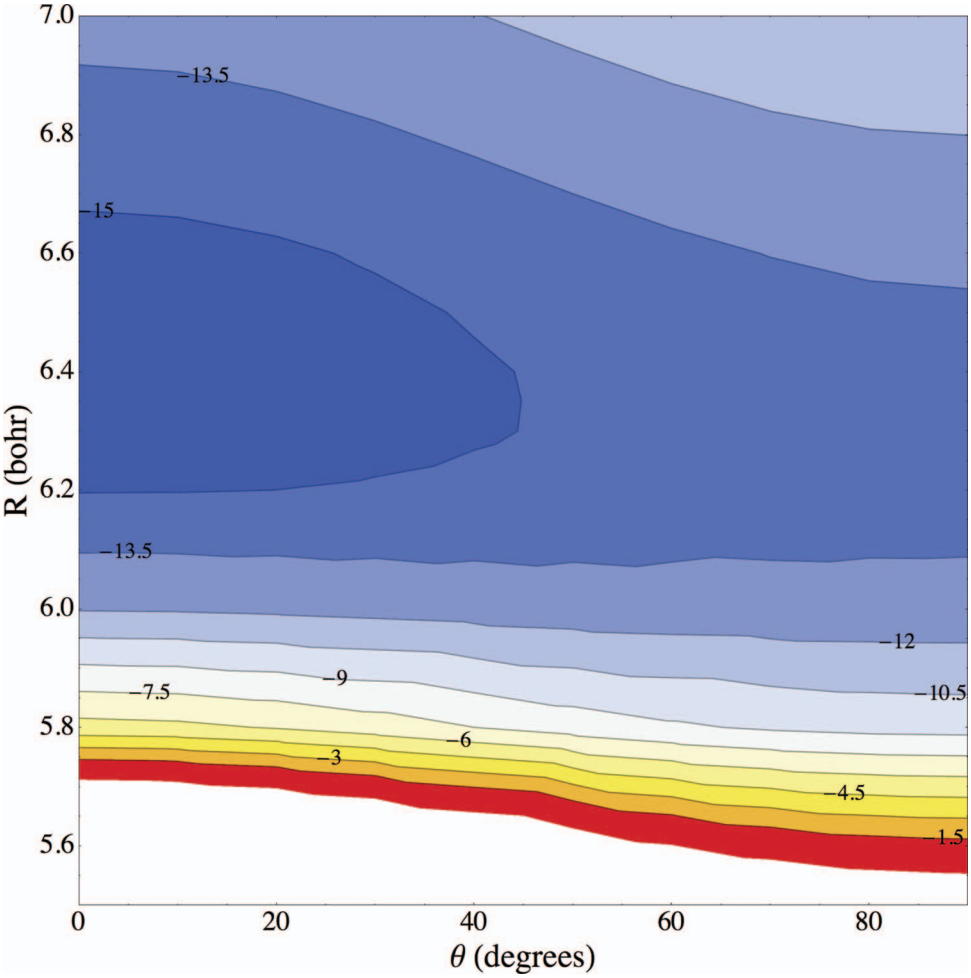

FIG. 1. Contour plot of the symmetry unique fragment of the He–H_2 two-body potential energy surface as a function of (R, θ) for $r_{HH} = 1.448736$ bohrs. The energy unit is 1 K.

The 36 parameters A_{kli}, the four short-range nonlinear parameters α_k, 18 long-range constants C_{nkl}, and the damping parameter d were determined in the fitting process. One should note that the interaction potential can be expected to exhibit a more complicated anisotropy at small R than at large R: therefore, the short-range part of $V(r_{HH}, R, \theta)$, the first line of Eq. (3), involves a longer expansion in Legendre polynomials than the long-range part. Moreover, an accurate calculation of the r_{HH}-dependent C_6 asymptotic constant[76] showed a nearly perfect linear dependence in the range of r_{HH} considered here. Thus, while a term quadratic in r_{HH} is somewhat helpful in fitting the long-range part of $V(r_{HH}, R, \theta)$, a cubic term is not needed at all; such a term is, however, absolutely crucial in the short-range part.

The fitting process was initiated by looking at all *ab initio* data with $R \geq 10$ bohrs and fitting the long-range constants C_{nkl} to these points only assuming no damping. Subsequently, the C_{nkl} values were frozen while the remaining parameters were fitted to all *ab initio* data points. The quantity that was minimized in the fitting process was the uncertainty-weighted root mean square error (RMSE)

$$\mathrm{RMSE}_\sigma = \sqrt{\frac{1}{n}\sum_{\text{data points}} \frac{(E_{\text{fit}} - E_{ab\ initio})^2}{\sigma^2}}, \qquad (5)$$

where σ is an estimated uncertainty of $E_{ab\ initio}$, computed in a manner that is a generalization of the estimates used for the two characteristic points (Sec. III)

$$\sigma = (\sigma_{\mathrm{CCSD(T)}} + \sigma_{\mathrm{CCSDT}}) \cdot 1.2, \qquad (6)$$

where the uncertainties of the (CBS-extrapolated) CCSD(T) and CCSDT terms are computed in the same way as in Sec. III and the factor of 1.2 is intended to account for the uncertainties of the FCI and DBOC terms and the neglect of the relativistic correction, generalizing the value of 0.010 K assumed for the characteristic points. The weighted RMSE defined in Eq. (5) favors a reproduction of different *ab initio* two-body energies to similar fractions of their uncertainties; we adopted this form as we want to assure that the fitted values fall into the $(E_{ab\ initio} - \sigma, E_{ab\ initio} + \sigma)$ intervals for all points. Our 59-parameter fit does possess this desirable property—the fitted energies deviate from the *ab initio* ones by 0.14σ on the average and 0.83σ at most. In terms of relative differences, the calculated and fitted interaction energies differ by 0.12% on the average and 46.1% at most, with all relative errors larger than 0.6% occurring at $R = 5.5 - 6.0$ bohrs, the distances closest to where the PES crosses zero. To conclude, our 3D fit is fully consistent with the *ab initio* uncertainties even as the latter might be considered somewhat too conservative.

TABLE IV. Parameters of the 2D and 3D potentials $V(R,\theta)$ and $V(r_{HH}, R, \theta)$ (Eqs. (7) and (3), respectively) fitted to the accurate two-body He–H_2 interaction energies computed in this work. The units of all parameters are such that Eqs. (7) and (3) give the potential in Kelvin when R and r_{HH} are in bohr.[39] Not all digits listed are significant: the extra digits are given for consistency with the programmed expressions. In the lower part of the table, the values of the asymptotic constants C_{nkl} are repeated in atomic units (hartree bohr^n) to facilitate comparisons with literature.

	2D	3D			
Parameter	k = 0	k = 0	k = 1	k = 2	k = 3
α_k	2.01155883	1.90211567	1.89629405	2.60059283	1.84852909
A_{k00}	1 377 925.52298519	3 334 383.60380158	3 609 379.04543074	9 157 918.37253898	− 731 643.11212337
A_{k01}	1 981 577.7039402	− 1 168 113.10369338	− 258 231.25352943	− 6 343 635.82181885	233 964.00717458
A_{k02}	− 266 546.8863106	95 343.59692526	− 22 623.45610335	1 144 649.72208008	− 15 008.61360829
A_{k03}	5904.56215244				
A_{k20}	− 99 350.22309416	− 2 656 122.74087459	4 327 331.07326328	5 347 011.18254587	2052.20960054
A_{k21}	854 663.16515132	847 915.1603025	− 1 376 067.74978694	− 3 931 124.07333934	120 048.72344083
A_{k22}	− 125 851.80420149	− 55 563.49555373	92 606.12637049	500 893.92781881	− 11 744.94220195
A_{k23}	4670.51556659				
A_{k40}	263 921.00275228	− 542 641.3156615	788 190.0995706	− 2 081 778.35428045	− 49 522.30546869
A_{k41}	− 158 978.07103669	217 509.80753925	− 289 088.46172645	1 109 111.674834	48 835.07041915
A_{k42}	41 806.80926751	− 19 306.06640687	25 035.56370694	− 260 552.10717593	− 5650.23395223
A_{k43}	− 3643.84776417				
d	3.16011458		3.97362189		
C_{6k0}	1.25504634×10^6	2.28959703×10^5	7.20151513×10^5	-8.36637311×10^3	
C_{6k2}	1.17102309×10^5	-3.48268888×10^4	4.61987518×10^4	4.04895328×10^4	
C_{8k0}	2.17129022×10^7	8.27533494×10^6	-4.21069871×10^6	9.32571992×10^6	
C_{8k2}	5.99393636×10^6	1.02403258×10^7	-1.96080228×10^7	1.15102734×10^7	
C_{10k0}	8.43596194×10^7	-3.30080231×10^8	7.66092689×10^8	-3.32557998×10^8	
C_{10k2}	1.26398210×10^8	-8.38176012×10^7	7.66008832×10^7	4.76561641×10^7	
C_{6k0} [a.u.]	3.974	0.725	2.281	− 0.026	
C_{6k2} [a.u.]	0.371	− 0.110	0.146	0.128	
C_{8k0} [a.u.]	68.76	26.21	− 13.33	29.53	
C_{8k2} [a.u.]	18.98	32.43	− 62.09	36.45	
C_{10k0} [a.u.]	267.2	− 1045.3	2426.1	− 1053.1	
C_{10k2} [a.u.]	400.3	− 265.4	242.6	150.9	

It is also worthwhile to construct a simplified, 2D potential $V(R,\theta)$ with the monomer flexibility effects neglected. A proper selection of monomer geometry that maximizes the predictive power of the rigid potential has been investigated quite extensively.[77,78] It has been shown that freezing the monomer bond length at the equilibrium value r_e (1.4015 bohrs for H_2[46]) is quite a poor choice if the resulting rigid potential is to be used to reproduce experimental quantities such as spectra. A much better approach is to use $\langle r_{HH}\rangle$, a bond length averaged over the vibrational function of the monomer. For the ground vibrational state of H_2, $\langle r_{HH}\rangle$ amounts to 1.448736 bohrs.[47,48] Therefore, we obtained a 2D analytical potential $V(R,\theta)$ by setting $r_{HH} = 1.448736$ bohrs and fitting a 20-parameter expression

$$V(R,\theta) = \sum_{l=0,2,4} (A_{0l0} + A_{0l1}R + A_{0l2}R^2 + A_{0l3}R^3) \times e^{-\alpha_0 R} P_l(\cos\theta) - \sum_{l=0,2} \sum_{n=6,8,10} f_n(dR)\frac{C_{n0l}}{R^n} P_l(\cos\theta) \quad (7)$$

to the 190 *ab initio* energies computed for this value of r_{HH} (the fitting parameters are analogous to Eq. (3) but only the r_{HH}^0 terms are present, thus the zero subscripts). The resulting potential recovers the *ab initio* data to 0.10σ or 0.05% on the average and 0.41σ or 1.0% in the worst case. A set of FORTRAN codes to compute our 2D and 3D potentials for a given geometry is provided in the supplementary material.[79]

The fitted parameters for both potentials are presented in Table IV. In addition, this table lists the values of the asymptotic constants C_{nkl} converted to atomic units. The constants for the 2D and 3D potentials are consistent at $r_{HH} = 1.448736$ bohrs, with the differences at the C_6, C_8, and C_{10} levels not exceeding 0.001, 0.12, and 8 a.u., respectively. Moreover, the C_{6k0} and C_{6k2} values computed from our 3D potential are in very good agreement with the *ab initio* r_{HH}-dependent asymptotic constants from the literature:[76,80] the differences with the values computed in Ref. 80 at $r_{HH} = 1.2$, 1.4, 1.449, and 1.65 bohrs do not exceed 0.07 a.u. The agreement is worse but still reasonable for C_{8k0} and C_{8k2} (with differences up to 20 a.u. or 30%), but the C_{10k0} values obtained from Table IV are significantly different from the *ab initio* ones of Ref. 80: as expected, higher asymptotic coefficients are progressively harder to recover by fitting to interaction energies at finite separations.

V. REPARAMETERIZATION OF THE BMP POTENTIAL

The 2D and 3D potentials constructed here, combined with Schwenke's fit of the H_2 monomer deformation energy,[46]

provide the most accurate representation of the He–H_2 vdW well to date. Thus, these potentials are well suited for computing quantities that are insensitive to higher-energy regions of the PES, such as bound state properties (Sec. VI), the second virial coefficient (Sec. VII), and elastic or rovibrationally inelastic scattering cross sections. However, our 2D and 3D potentials have not been designed to reproduce the strongly repulsive region of the PES (with $E_{\text{int, 2B}}$ above 3000 K) or the breaking of the H–H bond. On the other hand, the BMP potential[14] is suitable for all regions of the PES but, being based on fairly low-level *ab initio* data (by today's standards), is not nearly as accurate as our potential in the vdW region. Therefore, it is worthwhile to construct a potential that combines the strengths of our vdW-region fit and of the all-region BMP fit. To this end, we have generated a potential which we will refer to as BMPmod. It has the same functional form as the original BMP potential but has been fitted to our accurate *ab initio* data in the vdW region (and to the same data as the original BMP fit in all other regions).

The original BMP potential was fitted to 25 065 interaction energies: 16 703 points of the main grid, 3500 random conformations, 1887 "H–He" conformations representing a relatively close-together H–He pair far away from the remaining H atom, 830 "H–He+H" conformations with the H–H bond broken but both He–H interactions nonnegligible, and 2145 geometries in the vdW region with interaction energies computed previously or estimated from previous calculations. The latter set included 69 points computed by Tao[13] (with $r_{\text{HH}} = 1.28$, 1.449, and 1.618 bohrs), 374 "gen-Tao vdW" points (with $r_{\text{HH}} = 1.28$, 1.449, and 1.618 bohrs) representing estimated interaction energies at the Ref. 13 level, and 1702 points (with $r_{\text{HH}} \leq 1.1$ bohrs or $r_{\text{HH}} \geq 1.8$ bohrs) generated from the lower-level Schaefer-Köhler (SK)[11] and (slightly modified) MR[12] surfaces. Boothroyd *et al.*[14] computed 3500 additional random conformations that were not used in the fit to provide an independent validation of the fitted PES.

There are two slight inconsistencies between the formulas for the BMP potential published in Ref. 14 and the FORTRAN code that computes this potential, given in the supplementary material to that reference. First, the programmed formula for $F_H(R_a)$ ($F_H(R_b)$) is missing an overall factor of $1/R_a$ ($1/R_b$) compared to Eq. (9) in Ref. 14. Second, in the $n = 6$ term of Eq. (16) of Ref. 14, the coefficient $a_D(6) = -2.80333655$ hartree bohr^6 is inferred from the experimental H–He potential[81] and not optimized. Additionally, seven parameters (the ones in the expressions for $A_d(r)$ and $\alpha_d(r)$, Eqs. (13) and (14) of Ref. 14) were not fitted to the *ab initio* data points, but to accurate asymptotic values of Thakkar *et al.*[76] The remaining 104 "true" fit parameters include the "softening parameter of the switch function" that was fixed at 6.0 in Eqs. (17) and (18) of Ref. 14 but is considered an adjustable parameter (assumed to be the same in both equations). In our refitting process leading to the BMPmod potential, we will adhere to the programmed BMP expressions rather than the text of Ref. 14 to ensure consistency with the existing applications of the BMP potential.

Our accurate *ab initio* data cover a range of r_{HH} from 1.1 to 1.75 bohrs, encompassing all the Tao and "gen-Tao vdW" points. Therefore, we fitted the BMPmod potential to a set of 26 270 interaction energies: compared to the original BMP fitting set, our 1900 *ab initio* points were added while the 443 Tao and "gen-Tao vdW" points and 252 of the MR and SK points (those with $r_{\text{HH}} = 1.1$ bohrs) were removed. We fixed the seven $A_d(r)$ and $\alpha_d(r)$ parameters at their asymptotics-fitted BMP values and freely optimized the 104 "true" parameters including the "softening parameter." We employed the original BMP weights for all fit points other than ours. For our energies, we computed the weights in the same way as for the other vdW He–H_2 points in Ref. 14, that is, as products of four factors that depend on E_{int}, method, R, and r_{HH}, respectively. Taking into account the fact that our points are significantly more accurate than the vdW points in the original BMP fitting set, we assumed a method-dependent weighting factor of 5, compared to the values of 1 for the MR[12] and SK[11] points, 2 for the estimated "gen-Tao vdW" values, and 3 for the computed points of Tao.[13] Note that these weighting factors are *squared* in the definition of the weighted RMSE adopted in Ref. 14 and adhered to in this section. The weighted RMSE was minimized with respect to the 104 "true" parameters using a Powell optimization routine from the SCIPY package[82] and the original BMP parameters as a starting point. For consistency, we have verified that the original BMP parameters are close to optimal on the original BMP data set: specifically, a reoptimization of 104 "true" parameters only lowered the weighted RMSE on the 25 065-point data set from 1.383 millihartree (consistent with Table II of Ref. 14) to 1.374 millihartree. At the same time, a significantly larger change in parameters is needed to account for our accurate vdW-region results: on the 26 270-point BMPmod data set, the original BMP parameters lead to a weighted RMSE of 2.497 millihartree and a reoptimization reduces this RMSE to 1.416 millihartree.

Different measures of the accuracy of the BMP and BMPmod fits on several data sets are presented in Table V. The upper part of this table lists the same quantities as Table II in Ref. 14: the RMSE with full weights employed in the BMP and BMPmod fitting process and the RMSE with weights that depend on the total binding energy E (which includes the H–H binding energy) only

$$w_E(E) = \begin{cases} 1 & E \leq 0.2 \text{ hartree}, \\ (0.2 \text{ hartree})/E & E > 0.2 \text{ hartree}. \end{cases} \quad (8)$$

As the values in Table V show, the change of RMSE on the original BMP fitting set between the BMP and BMPmod fits is minor. This can be viewed as a confirmation that the refitting process has not adversely affected the accuracy of the fit outside the vdW well. It is also encouraging that the BMP potential expression is able to fit the BMPmod set including our accurate vdW-region data to practically the same RMSE as for the original BMP set. Thus, the refitted BMPmod potential maintains the overall accuracy of BMP across the entire PES. This statement is further supported by the BMP and BMPmod potentials reproducing the additional 3500 random configurations computed in Ref. 14 to virtually the same accuracy (note that these configurations were not included in the fitting set for either potential).

TABLE V. Statistical measures of the accuracy of the original BMP[14] and BMPmod potentials with respect to different *ab initio* results. The upper part of the table pertains to total binding energies $V_{tot} = E_{int,2B} + V_{H-H}$ (where the last term, the binding energy of the H_2 molecule, is computed using the potential of Schwenke[46]), and lists the full-weighted RMSE and energy-only-weighted RMSE (in millihartree) in accordance with Table II of Ref. 14. The lower part of the table shows mean unsigned errors (MUE), mean unsigned relative errors (MURE), and mean unsigned errors relative to the uncertainty (MUEσ) for the $E_{int,2B}$ term alone using different sets of data in the vdW He–H_2 region. For this part, the accuracy of our 3D potential, fitted to the *ab initio* data computed in this work, is shown for comparison.

Quantity		BMP[14]	BMPmod	Our 3D
Original BMP fitting set (25 065 points)				
Full-weighted RMSE	(millihartree)	1.383	1.433	
Energy-weighted RMSE	(millihartree)	1.320	1.328	
BMPmod fitting set (26 270 points)				
Full-weighted RMSE	(millihartree)	2.497	1.416	
Energy-weighted RMSE	(millihartree)	1.289	1.297	
Unfitted random BMP configurations (3500 points)				
Full-weighted RMSE	(millihartree)	1.413	1.411	
Energy-weighted RMSE	(millihartree)	1.400	1.398	
Tao[13] and "gen-Tao vdW" set (443 points)				
MUE	(Kelvin)	7.53	8.57	22.4
MURE	(%)	6.30	8.18	9.28
Max. URE	(%)	147.6	213.5	303.1
MR[12] and SK[11] points, $r_{HH} \leq 0.9$ or ≥ 1.8 bohrs (1450 points)				
MUE	(Kelvin)	10.8	10.8	10.7
MURE	(%)	66.4	64.6	48.6
Max. URE	(%)	15 070	14 781	7231
Our *ab initio* set (1900 points)				
MUE	(Kelvin)	5.88	3.87	0.10
MURE	(%)	12.9	4.31	0.12
Max. URE	(%)	4611	1543	46.1
MUEσ	(σ units)	11.4	4.93	0.14
Max. UEσ	(σ units)	38.0	45.0	0.83

A reliable reproduction of the *interaction energies* in the vdW region is a more ambitious goal than a reliable reproduction of the *total binding energies* (relative to isolated H, H, and He atoms) outside of the vdW well. Indeed, the two-body vdW interaction energy is only a small fraction of the total binding energy. Moreover, our new interaction energies in the vdW region are substantially more accurate than the energies in the original BMP fitting set.[14] Thus, as the lower part of Table V illustrates, the relative errors of the original BMP potential on our accurate two-body vdW interaction energies are rather high, 13% (or 11σ, where σ is the estimated uncertainty of the *ab initio* result) on the average. The BMPmod reparameterization brings the errors down to 4.3% or 4.9σ on the average, better but still far from perfect. Therefore, we have included in the lower part of Table V the results for our own 3D potential, fitted to the 1900 *ab initio* points as described in Sec. IV. One can clearly see that the vdW-region accuracy afforded by our 3D fit is unattainable for a full-PES potential such as BMP or BMPmod and it is advantageous to have both potentials available for different applications. Understandably, the original BMP potential is the best performer on the Tao[13] and "gen-Tao vdW" sets, now superseded by our new *ab initio* data to which the other two potentials were fitted. For the MR[12] and SK[11] data with the r_{HH} distance further from the equilibrium ($\leq$0.9 bohr or $\geq$1.8 bohrs), no potential provides a good accuracy of interaction energies. Interestingly, our 3D vdW-region fit actually describes the MR and SK points slightly better than either BMP or BMPmod despite the fact that it was not fitted to any points with comparable values of r_{HH}. Overall, the results of Table V confirm that the goals of improving the accuracy of the vdW region and maintaining the accuracy of the original BMP potential[14] for other regions of the PES have been fulfilled by our BMPmod reparameterization.

VI. THE HE–H_2 BOUND STATE

It is well known that the He–H_2 van der Waals complex, similar to the helium dimer, is an example of a quantum halo system[83] where the binding is so weak that the bound-state wavefunction extends into very large intermolecular distances. Specifically, the complexes of ^{4}He and ^{3}He with *para*-H_2 exhibit a single bound state with a dissociation energy not exceeding 0.1 K (the complexes with *ortho*-H_2 are unbound due to the rotational zero-point energy). The properties of the lone bound state have been computed by several authors[15–17] using the BMP potential[14] (an older work[84] utilized the MR potential[12]). However, the energy E of the bound state and the average intermolecular distance $\langle R \rangle$ are highly sensitive to the details of the interaction potential. The errors of the BMP potential in the vdW region (BMP underestimates the vdW well depth by about 1 K) are likely large enough to produce a

TABLE VI. Properties of the lone bound state of ^{4}He–H_2 and ^{3}He–H_2: the energy E and the average distance $\langle R\rangle$ and $\sqrt{\langle R^2\rangle}$,[39] computed in this work and in previous studies. Our calculations employing the BMP and BMPmod potentials were restricted to 2D by setting $r_{HH} = 1.448736$ bohrs.

		^{4}He–H_2			^{3}He–H_2		
Ref.	Potential	E (cm^{-1})	$\langle R\rangle$ (Å)	$\sqrt{\langle R^2\rangle}$ (Å)	E (cm^{-1})	$\langle R\rangle$ (Å)	$\sqrt{\langle R^2\rangle}$ (Å)
This work	Our 2D	− 0.06657	10.91	12.94	− 0.01489	19.42	24.65
This work	BMP[14]	− 0.03812	13.20	16.07	− 0.00388	32.98	41.84
This work	BMPmod	− 0.05993	11.28	13.45	− 0.01169	21.35	27.26
84	MR[12]	− 0.0298			− 0.0016		
15	BMP[14]	− 0.03634	13.42		− 0.02916[a]	28.55	
16	BMP[14]	− 0.03640			− 0.00327		
17	BMP[14]	− 0.0368	13.37	16.31	− 0.00348	32.92	41.59

[a]Most likely a misprint in Ref. 15: a value of −0.002916 would be very reasonable.

substantial error in the bound state energy. It is thus worthwhile to investigate the He–H_2 bound-state properties using our newly constructed potential.

A formally exact calculation of the bound-state energy requires solving the nuclear Schrödinger equation in all three dimensions including r_{HH}. It is, however, much simpler to solve for the bound state using a rigid-monomer $V(R,\theta)$ potential, and such an approximation can be expected to give very reasonable results when r_{HH} is fixed at its vibrationally averaged value.[78] We will follow this simplified approach and locate the bound state using the algorithms and software developed by Hutson.[85,86] Specifically, a system of coupled differential equations corresponding to different rotational quantum numbers is propagated radially (using the diabatic modified log-derivative method of Manolopoulos[87]) outwards and inwards to a common matching point in the classically allowed region. Then, a search for zero eigenvalues of the matching matrix is used to locate the bound-state energy. While the wavefunction of the complex is not computed, expectation values such as $\langle R\rangle$ can be obtained by finite-field techniques.[88] The results are highly sensitive to the reduced mass of the complex. The masses assumed here, 1.34056296391 and 1.20819695125 amu for ^{4}He–$^1H_2$ and ^{3}He–$^1H_2$, respectively, were obtained from the *atomic* masses of the appropriate isotopes. The interaction potential $V(R,\theta)$ is supplied as an expansion in even-l Legendre polynomials $P_l(\cos\theta)$. This expansion terminates after P_4 for our 2D and 3D potentials; for the BMP[14] and BMPmod potentials studied for comparison (with r_{HH} fixed at 1.448736 bohrs), the terms through P_{10} were taken into account.

The energies E and expectation values $\langle R\rangle$ and $\sqrt{\langle R^2\rangle}$ of the ^{4}He–H_2 and ^{3}He–H_2 bound states obtained using our 2D potential, BMP, and BMPmod are presented in Table VI and compared to the literature data. The values that we computed using the original BMP potential[14] are in good agreement with previous BMP-based calculations,[15–17] indicating that our rigid-H_2 approximation is reliable (note, however, that other calculations also employed some approximations and may have used slightly different reduced masses). However, the differences between the bound state properties predicted by BMP and by the potentials constructed here are quite dramatic. An error of about 1 K in the BMP well depth (for $r_{HH} = 1.448736$ bohrs) translates into an underestimation of the bound state energy by a factor of about 1.7 for ^{4}He–H_2 and nearly 4 for ^{3}He–H_2. The BMPmod potential (whose error in the well depth is about 0.05 K) performs much better although the bound-state energies are still somewhat underestimated. Overall, the results in Table VI illustrate the ultra-high demands imposed on the potential accuracy by a very weakly bound state calculation. Similar, or even stricter, demands have been observed for the ^{4}He–^{4}He dimer.[89] Under these circumstances, a dedicated vdW-region potential fitted to accurate *ab initio* data, such as our 2D potential of Sec. IV, should be recommended for computing the bound state properties.

VII. SECOND VIRIAL COEFFICIENT

The knowledge of an analytic He–H_2 potential enables one to calculate a large variety of bulk properties of the hydrogen-helium mixture such as the interaction second virial coefficient, the binary diffusion coefficient, the viscosity coefficient, and several others. A multitude of experimental data for these properties enables an extensive comparison that benefits both theory (by helping quantify the residual errors of the theoretical approaches) and experiment (by providing reasonable baselines for future measurements and, in cases of extremely high accuracy, enabling one to establish new, improved thermophysical standards[90]). For the He–H_2 complex, a large variety of transport and relaxation properties has been computed *ab initio*[91] using the interaction potentials of Tao[13] and Schaefer and Köhler,[11] that is, the data used to anchor the vdW-region behavior of the BMP potential.[14] While an extensive calculation of a large number of such properties is beyond the scope of the present work, we employed our accurate analytic 2D potential to compute the interaction second virial coefficient $B_{He-H_2}(T)$ for a range of temperatures T. The second virial coefficient was approximated by a sum of its classical value $B^{(0)}(T)$ and the first quantum correction $B^{(1)}(T)$. The latter contribution is a sum of three terms representing the radial, angular, and Coriolis corrections.[92,93] All contributions were numerically integrated using a program written by Heijmen[93,94] interfaced to our 2D potential routine. The reduced mass of the dimer was set to 1.33333 amu, corresponding to the ^{4}He–$^1H_2$ isotopomer of the complex.

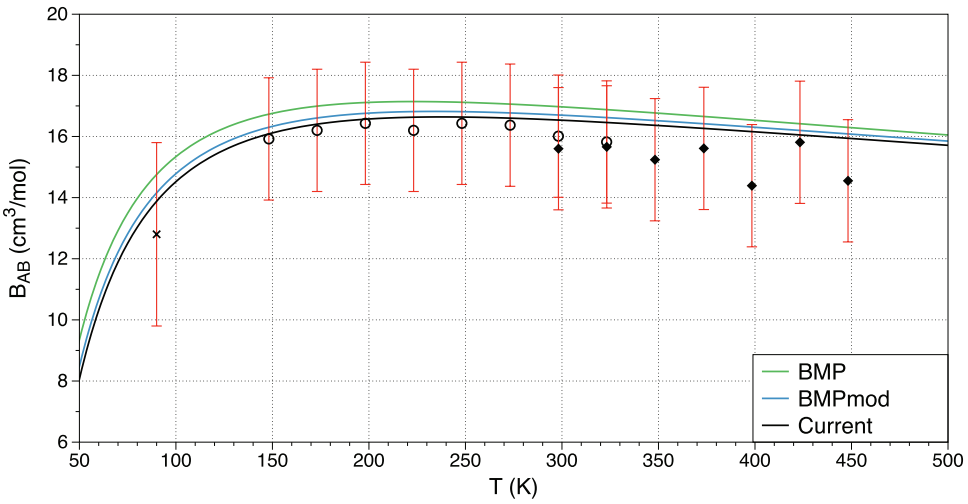

FIG. 2. The He–H_2 interaction second virial coefficient as a function of temperature. The results with our 2D potential and with the original BMP[14] and BMPmod potentials (truncated to $r_{HH} = 1.448736$ bohrs) are presented. The experimental data are taken from Refs. 97 (x), 98 (circles), and 96 (diamonds).

The rotational constant of H_2 was set to 60.853 cm^{-1}.[95] The second virial coefficients predicted by the BMP[14] and BMPmod potentials (with r_{HH} set to 1.448736 bohrs) were also computed for comparison. The resulting $B(T)$ curves are displayed in Fig. 2 which also contains a number of experimental values[96–98] with the uncertainties estimated in Ref. 91.

The approximation of the second virial coefficient by its classical value plus the first quantum correction breaks down for temperatures below 50 K as indicated by the radial term of $B^{(1)}(T)$ dominating the entire $B(T)$. While a complete formula for the second quantum correction $B^{(2)}(T)$ for an atom-linear molecule complex is, to our knowledge, not available, we tried approximating $B^{(2)}(T)$ by its isotropic part using the known atom-atom expression.[99,100] The $B^{(2)}(T)$ correction amounts to -1.0 cm^3 mol^{-1} at 50 K and progressively becomes more negative as the temperature decreases. While it is likely that an inclusion of the approximate $B^{(2)}(T)$ term would improve the accuracy of the second virial coefficient in the 25–50 K range, there are no experimental data to confirm it, and neither $B^{(0)} + B^{(1)}$ nor $B^{(0)} + B^{(1)} + B^{(2)}$ agree with experiment below 25 K. For such low temperatures, a full inclusion of quantum effects, for example, through a path-integral Monte Carlo simulation,[44] would be necessary to obtain reliable results. However, for the 50–500 K range displayed in Fig. 2 the present treatment is fully adequate. While the differences between our potential, BMP, and BMPmod are relatively minor and all computed virial coefficients are within experimental error bars, our potential gives the best agreement with the actual measured values. It should be noted that all three potentials give $B(T)$ values that are lower (and more accurate) than the SK potential[11] employed in Ref. 91 which is at, or slightly beyond, the upper limit of the experimental error bars.

VIII. SUMMARY

We have obtained a vastly improved interaction potential for the He–H_2 van der Waals complex using a variety of state-of-the-art electronic structure methods (CCSD(T), CCSDT, FCI) and large correlation-consistent basis sets in conjunction with midbond functions and CBS extrapolations. While the relativistic effects on the interaction potential turned out to be negligible, the inclusion of the diagonal Born-Oppenheimer correction was necessary. At the near-minimum (r_{HH}, R, θ) = (1.448736 bohrs, 6.4 bohrs, 0°) geometry, our best result for the two-body He–H_2 interaction energy amounts to -15.864 ± 0.035 K while the result using the "production-level" bases employed for all data points is -15.870 ± 0.065 K. These values should be compared with the previous best result of -14.90 K.[13]

The *ab initio* calculations have been performed for 1900 symmetry-unique He–H_2 geometries corresponding to the H–H bond length ranging from 1.1 to 1.75 bohrs. The resulting interaction energies have provided data for fitting a 59-parameter analytical 3D potential. We have also fitted a 20-parameter rigid-monomer (2D) potential and reoptimized 104 parameters in the full-PES potential of Boothroyd, Martin, and Peterson[14] to better reproduce our new accurate data in the van der Waals region.

The *ab initio* He–H_2 potentials constructed in this work can be used to compute various spectroscopic and bulk properties of the hydrogen-helium mixture that can be compared to experimental data. We made the first step in this direction and computed the lone bound vibrational state of ^{4}He–H_2 and ^{3}He–H_2 and the interaction second virial coefficient of the mixture. However, our new potentials are likely to provide accurate values of many other properties such as the diffusion, viscosity, and thermal conductivity coefficients[91] and rovibrationally inelastic scattering cross sections.[19] Moreover, as highly accurate He–He and H_2–H_2 interaction potentials have been available for a while,[42,101] this work eliminates the weakest link in an *ab initio* description of hydrogen-helium clusters, systems of broad interest ranging from spectroscopy to low-temperature physics to astrophysics.

ACKNOWLEDGMENTS

This work was supported by startup funding from Auburn University. We thank Dr. Arnold Boothroyd for granting us

the permission to provide a modification of his BMP potential code in the supplementary material,[79] and Dr. Jim Mehl for the code to calculate the isotropic part of the second quantum correction to the second virial coefficient. We also thank Dr. Wojciech Cencek and Dr. Allan Harvey for helpful discussions.

[1] T. Abel, G. L. Bryan, and M. L. Norman, Science **295**, 93 (2002).
[2] S. C. O. Glover and T. Abel, Mon. Not. R. Astron. Soc. **388**, 1627 (2008).
[3] H. C. Harris, E. Gates, G. Gyuk, M. Subbarao, S. F. Anderson, P. B. Hall, J. A. Munn, J. Liebert, G. R. Knapp, D. Bizyaev, E. Malanushenko, V. Malanushenko, K. Pan, D. P. Schneider, and J. A. Smith, Astrophys. J. **679**, 697 (2008).
[4] S. Grebenev, B. Sartakov, J. P. Toennies, and A. F. Vilesov, Science **289**, 1532 (2000).
[5] F. Mezzacapo and M. Boninsegni, Phys. Rev. Lett. **100**, 145301 (2008).
[6] P. Barletta, J. Tennyson, and P. F. Barker, Phys. Rev. A **78**, 052707 (2008).
[7] C. W. Wilson, Jr., R. Kapral, and G. Burns, Chem. Phys. Lett. **24**, 488 (1974).
[8] P. J. M. Geurts, P. E. S. Wormer, and A. van der Avoird, Chem. Phys. Lett. **35**, 444 (1975).
[9] W. Meyer, P. C. Hariharan, and W. Kutzelnigg, J. Chem. Phys. **73**, 1880 (1980).
[10] U. E. Senff and P. G. Burton, J. Phys. Chem. **89**, 797 (1985).
[11] J. Schaefer and W. E. Köhler, Physica A **129**, 469 (1985).
[12] P. Muchnick and A. Russek, J. Chem. Phys. **100**, 4336 (1994).
[13] F.-M. Tao, J. Chem. Phys. **100**, 4947 (1994).
[14] A. I. Boothroyd, P. G. Martin, and M. R. Peterson, J. Chem. Phys. **119**, 3187 (2003).
[15] F. A. Gianturco, T. González-Lezana, G. Delgado-Barrio, and P. Villareal, J. Chem. Phys. **122**, 084308 (2005).
[16] Y. Xiao and B. Poirier, J. Phys. Chem. A **110**, 5475 (2006).
[17] P. Barletta, Eur. Phys. J. D **53**, 33 (2009).
[18] H. Suno, J. Chem. Phys. **132**, 224311 (2010).
[19] T.-G. Lee, C. Rochow, R. Martin, T. K. Clark, R. C. Forrey, N. Balakrishnan, P. C. Stancil, D. R. Schultz, A. Dalgarno, and G. J. Ferland, J. Chem. Phys. **122**, 024307 (2005).
[20] M. Audibert, R. Vilaseca, J. Lukasik, and J. Ducuing, Chem. Phys. Lett. **37**, 408 (1976).
[21] J. L. Nolte, B. H. Yang, P. C. Stancil, T.-G. Lee, N. Balakrishnan, R. C. Forrey, and A. Dalgarno, Phys. Rev. A **81**, 014701 (2010).
[22] J. L. Nolte, P. C. Stancil, T.-G. Lee, N. Balakrishnan, and R. C. Forrey, Astrophys. J. **744**, 62 (2012).
[23] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[24] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[25] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).
[26] F.-M. Tao and Y.-K. Pan, J. Phys. Chem. **95**, 3582 (1991).
[27] C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Chem. Rev. **112**, 4 (2012).
[28] L. Kong, F. A. Bischoff, and E. F. Valeev, Chem. Rev. **112**, 75 (2012).
[29] K. E. Riley and P. Hobza, WIREs Comput. Mol. Sci. **1**, 3 (2011).
[30] E. G. Hohenstein and C. D. Sherrill, WIREs Comput. Mol. Sci. **2**, 304 (2012).
[31] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 9221 (1998).
[32] Y. J. Bomble, J. F. Stanton, M. Kállay, and J. Gauss, J. Chem. Phys. **123**, 054101 (2005).
[33] R. Hellmann, E. Bich, and E. Vogel, Mol. Phys. **106**, 133 (2008).
[34] K. Patkowski and K. Szalewicz, J. Chem. Phys. **133**, 094304 (2010).
[35] P. Jankowski, A. R. W. McKellar, and K. Szalewicz, Science **336**, 1147 (2012).
[36] J. R. Lane, J. Chem. Theory Comput. **9**, 316 (2013).
[37] J. Řezáč, L. Šimová, and P. Hobza, J. Chem. Theory Comput. **9**, 364 (2013).
[38] J. Řezáč and P. Hobza, J. Chem. Theory Comput. **9**, 2151 (2013).
[39] The conversion factors employed were 1 hartree = 315 774.65 K and 1 K = 0.6950356 cm^{-1} for the energy and 1 bohr = 0.529177209 Å for the distance.
[40] W. Cencek and K. Szalewicz, Int. J. Quantum Chem. **108**, 2191 (2008).
[41] W. Cencek, M. Przybytek, J. Komasa, J. B. Mehl, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **136**, 224303 (2012).
[42] K. Patkowski, W. Cencek, P. Jankowski, K. Szalewicz, J. B. Mehl, G. Garberoglio, and A. H. Harvey, J. Chem. Phys. **129**, 094304 (2008).
[43] R. J. Hinde, J. Chem. Phys. **128**, 154308 (2008).
[44] G. Garberoglio, P. Jankowski, K. Szalewicz, and A. H. Harvey, J. Chem. Phys. **137**, 154308 (2012).
[45] X. Li, A. Mandal, E. Miliordos, and K. L. C. Hunt, J. Chem. Phys. **136**, 044320 (2012).
[46] D. W. Schwenke, J. Chem. Phys. **89**, 2076 (1988).
[47] R. J. LeRoy and J. M. Hutson, J. Chem. Phys. **86**, 837 (1987).
[48] P. Jankowski and K. Szalewicz, J. Chem. Phys. **123**, 104301 (2005).
[49] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **100**, 2975 (1994).
[50] S. L. Mielke, B. C. Garrett, and K. A. Peterson, J. Chem. Phys. **116**, 4142 (2002).
[51] R. J. Gdanitz, J. Chem. Phys. **113**, 5145 (2000).
[52] R. Podeszwa, R. Bukowski, and K. Szalewicz, J. Phys. Chem. A **110**, 10345 (2006).
[53] R. Podeszwa, K. Patkowski, and K. Szalewicz, Phys. Chem. Chem. Phys. **12**, 5974 (2010).
[54] H. L. Williams, E. M. Mas, K. Szalewicz, and B. Jeziorski, J. Chem. Phys. **103**, 7374 (1995).
[55] M. Jeziorska, W. Cencek, K. Patkowski, B. Jeziorski, and K. Szalewicz, Int. J. Quantum Chem. **108**, 2053 (2008).
[56] K. Patkowski, J. Chem. Phys. **137**, 034103 (2012).
[57] M. Douglas and N. M. Kroll, Ann. Phys. **82**, 89 (1974).
[58] B. A. Hess, Phys. Rev. A **33**, 3742 (1986).
[59] J. Gauss, A. Tajti, M. Kállay, J. F. Stanton, and P. G. Szalay, J. Chem. Phys. **125**, 144111 (2006).
[60] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2010.1, a package of *ab initio* programs, 2010, see http://www.molpro.net.
[61] J. Stanton, J. Gauss, M. Harding, P. Szalay, A. Auer, R. Bartlett, U. Benedikt, C. Berger, D. Bernholdt, Y. Bomble, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W. Lauderdale, D. Matthews, T. Metzroth, D. O'Neill, D. Price, E. Prochnow, K. Ruud, F. Schiffmann, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, and J. Watts, CFOUR, a quantum chemical program package, containing the integral packages MOLECULE (J. Almlöf and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de (accessed 6 November 2012).
[62] J. Olsen, *LUCIA, A full CI, general active space program*, with contributions from H. Larsen and M. Fulscher, 2000.
[63] DALTON, a molecular electronic structure program, Release 2.0, 2005, see http://daltonprogram.org.
[64] T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[65] G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[66] C. Hättig, D. P. Tew, and A. Köhn, J. Chem. Phys. **132**, 231102 (2010).
[67] K. Patkowski, J. Chem. Phys. **138**, 154101 (2013).
[68] J. Mitroy, S. Bubin, W. Horiuchi, Y. Suzuki, L. Adamowicz, W. Cencek, K. Szalewicz, J. Komasa, D. Blume, and K. Varga, Rev. Mod. Phys. **85**, 693 (2013).
[69] W. Tung, M. Pavanello, and L. Adamowicz, J. Chem. Phys. **133**, 124106 (2010).
[70] L. Wolniewicz, J. Chem. Phys. **99**, 1851 (1993).
[71] J. S. Sims and S. A. Hagstrom, J. Chem. Phys. **124**, 094101 (2006).
[72] H. Nakatsuji, H. Nakashima, Y. Kurokawa, and A. Ishikawa, Phys. Rev. Lett. **99**, 240402 (2007).
[73] K. Piszczatowski, G. Łach, M. Przybytek, J. Komasa, K. Pachucki, and B. Jeziorski, J. Chem. Theory Comput. **5**, 3039 (2009).
[74] K. Pachucki, Phys. Rev. A **82**, 032509 (2010).
[75] K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).
[76] A. J. Thakkar, Z.-M. Hu, C. E. Chuaqui, J. S. Carley, and R. J. LeRoy, Theor. Chim. Acta **82**, 57 (1992).
[77] E. M. Mas and K. Szalewicz, J. Chem. Phys. **104**, 7606 (1996).
[78] M. Jeziorska, P. Jankowski, K. Szalewicz, and B. Jeziorski, J. Chem. Phys. **113**, 2957 (2000).
[79] See supplementary material at http://dx.doi.org/10.1063/1.4824299 for the complete set of *ab initio* data and the FORTRAN codes calculating our 2D, 3D, and BMPmod potentials.
[80] P. E. S. Wormer, H. Hettema, and A. J. Thakkar, J. Chem. Phys. **98**, 7140 (1993).
[81] R. Gengenbach, C. Hahn, and J. P. Toennies, Phys. Rev. A **7**, 98 (1973).

[82]E. Jones, T. Oliphant, and P. Peterson, *SciPy: Open Source Scientific Tools for Python*, 2001, see http://www.scipy.org/.
[83]A. S. Jensen, K. Riisager, and D. V. Fedorov, Rev. Mod. Phys. **76**, 215 (2004).
[84]N. Balakrishnan, R. C. Forrey, and A. Dalgarno, Phys. Rev. Lett. **80**, 3224 (1998).
[85]J. M. Hutson, Comput. Phys. Commun. **84**, 1 (1994).
[86]J. M. Hutson, BOUND computer code, version 5, distributed by Collaborative Computational Project No. 6 of the UK Science and Engineering Research Council, 1993.
[87]D. E. Manolopoulos, Ph.D. thesis, University of Cambridge, 1988.
[88]J. M. Hutson, Chem. Phys. Lett. **151**, 565 (1988).
[89]M. Przybytek, W. Cencek, J. Komasa, G. Łach, B. Jeziorski, and K. Szalewicz, Phys. Rev. Lett. **104**, 183003 (2010).
[90]J. J. Hurly and M. R. Moldover, J. Res. Natl. Inst. Stand. Technol. **105**, 667 (2000).
[91]F. R. W. McCourt, D. Weir, G. B. Clark, and M. Thachuk, Mol. Phys. **103**, 17 (2005).
[92]R. T. Pack, J. Chem. Phys. **78**, 7217 (1983).
[93]R. Moszynski, T. Korona, T. G. A. Heijmen, P. E. S. Wormer, A. van der Avoird, and B. Schramm, Pol. J. Chem. **72**, 1479 (1998).
[94]T. G. A. Heijmen, T. Korona, R. Moszynski, P. E. S. Wormer, and A. van der Avoird, J. Chem. Phys. **107**, 902 (1997).
[95]K. P. Huber and G. Herzberg, "Constants of diatomic molecules," (data prepared by J. W. Gallagher and R. D. Johnson III) in *NIST Chemistry WebBook*, NIST Standard Reference Database No. 69, edited by P. J. Linstrom and W. G. Mallard (National Institute of Standards and Technology, Gaithersburg, MD, 2005), see http://webbook.nist.gov.
[96]C. W. Gibby, C. C. Tanner, and I. Masson, Proc. R. Soc. London, Ser. A **122**, 283 (1929).
[97]C. M. Knobler, J. J. M. Beenakker, and H. F. P. Knaap, Physica **25**, 909 (1959).
[98]J. Brewer and G. W. Vaughn, J. Chem. Phys. **50**, 2960 (1969).
[99]T. Kihara, Y. Midzuno, and T. Shizume, J. Phys. Soc. Jpn. **10**, 249 (1955).
[100]A. J. Thakkar, Mol. Phys. **36**, 887 (1978).
[101]M. Jeziorska, W. Cencek, K. Patkowski, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **127**, 124303 (2007).