Communication: Fitting potential energy surfaces with fundamental invariant neural network

Kejie Shao, Jun Chen, Zhiqiang Zhao, and Dong H. Zhang

Communication: Fitting potential energy surfaces with fundamental invariant neural network

Kejie Shao, Jun Chen, Zhiqiang Zhao, and Dong H. Zhang[a)]
State Key Laboratory of Molecular Reaction Dynamics and Center for Theoretical Computational Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, People's Republic of China and University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China.

A more flexible neural network (NN) method using the fundamental invariants (FIs) as the input vector is proposed in the construction of potential energy surfaces for molecular systems involving identical atoms. Mathematically, FIs finitely generate the permutation invariant polynomial (PIP) ring. In combination with NN, fundamental invariant neural network (FI-NN) can approximate any function to arbitrary accuracy. Because FI-NN minimizes the size of input permutation invariant polynomials, it can efficiently reduce the evaluation time of potential energy, in particular for polyatomic systems. In this work, we provide the FIs for all possible molecular systems up to five atoms. Potential energy surfaces for OH_3 and CH_4 were constructed with FI-NN, with the accuracy confirmed by full-dimensional quantum dynamic scattering and bound state calculations. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4961454]

Potential energy surfaces (PESs) play an important role in the theoretical study of molecular systems. With an accurately constructed PES, one can calculate rovibrational energy levels of a molecular system, or use quasiclassical trajectory (QCT), semiclassical approach, or quantum dynamics methods such as time-dependent wave packet method or multiconfiguration time-dependent Hartree method to investigate the dynamics of a molecular system. Splines and nonlinear fitting based on many-body expansions, functional forms, etc.,[1,2] are good choices in the construction of PES for systems containing no more than three atoms but are not of general use for larger systems. For molecular systems containing more than three atoms, some general fitting approaches have been developed. The most outstanding methods include permutation invariant polynomials (PIPs),[3–5] neural network (NN),[6–11] and permutation invariant polynomial neural network (PIP-NN).[12–14]

For a molecular system containing identical atoms, the potential energy should be permutation invariant with the permutations of like atoms. Therefore, the permutation symmetry is an important factor which needs to be considered in the construction of PESs.[3,10] With non-symmetric coordinates, the permutation invariance can be introduced into a PES by fitting the PES in a reduced non-symmetric zone and then symmetrizing the other geometries to this zone.[11,15] However, the gradient of the potential energy may not be derivable at the high symmetry points with this approach. Another way is by symmetrizing the data to construct permutation invariant neurons in the first hidden layer in NN, as used to construct the PES of H_3^+ by Prudente and his co-workers.[16] Besides these approaches, the permutation symmetry can be incorporated into the PES by using a permutation invariant basis. Gassner *et al.* used a symmetrized coordinate as the input vector of neural network.[17] The PIP method uses primary and secondary invariants as the fitting basis to generate permutation invariant polynomials while the coefficients are obtained by linear least squares fitting.[3] Behler *et al.* have developed a method for high-dimensional PESs and the permutation symmetry of the system is introduced by using symmetry functions.[10,18,19]

The extra requirement of permutation symmetry in PES leads to the extra addition of functional basis used for fitting. In other words, the dimension of vector space with the transformed symmetrized basis will be equal or larger than the original vector space. Only with a set of larger symmetrized basis it is possible to have a one to one mapping between the two bases. Bowman and co-workers developed the PIP method based on the theory of polynomial invariants and they have pointed out that it is a mathematical theorem: the primary and secondary invariants generate an invariant polynomial ring.[3] Guo and co-workers were recently inspired by PIP and proposed the PIP-NN fitting method.[12,13] In PIP-NN, the input vector of the neural network is replaced by a set of permutation invariant polynomials which contains all the polynomials truncated by a given degree. The degree for the truncation is the highest degree of the primary and secondary invariants. However, the number of polynomials increases nonlinearly with the degree bound. For the simple A_3B_2 system, for example, the highest degree of secondary invariants is 11 and there are 14 984 invariant polynomials with this degree bound. It is impractical to include all the invariant polynomials with the highest degree bound of secondary invariants for the system in the input. With a lower degree bound, say six, the number of polynomials can be reduced to 525, but more than half-secondary polynomials will be excluded.

---

a)Electronic mail: zhangdh@dicp.ac.cn

In this work, a new set of invariant polynomials (or called as invariants), known as fundamental invariants (FIs), is used as the input vector of neural network. Different to the primary and secondary polynomials in PIP and the polynomials in PIP-NN method, FI contains the least number of invariants which can generate all the invariant polynomials. We call the corresponding neural network with FI as the input vector as fundamental invariant neural network (FI-NN). It is worthwhile to point out that the mathematical background of fundamental invariants was well established decades ago.[20] Recently, Opalka and Domcke employed FI to generate invariant polynomials in their linear invariant polynomial fitting of multi-sheeted PESs for a XY_4 molecule.[21]

For a molecular system $A_iB_j \cdots X_p$, internuclear distances $(r_1, r_2, \ldots, r_n)$ are chosen as the variables set of the potential energy surfaces in order to deal with invariant properties under permutations of identical atoms. The permutation operator can be expressed as $\hat{g}$, which is an element of the direct product symmetric groups $G = S_i \times S_j \times \cdots \times S_p$, acting on the internuclear distances, here S_n is the symmetric group of degree n.

Let R be a set which contains all the polynomials in n variables $r = (r_1, r_2, \ldots, r_n)$ with real coefficients. R is a polynomial ring when it is an abelian group under addition and multiplication. The multiplication is left and right distributive with respect to addition. All the functions of $r = (r_1, r_2, \ldots, r_n)$ are elements of ring R. Invariant ring R^G is a subring of R, whose elements are invariant under the permutation operator of $\hat{g}$. R^G is finitely generated and has a minimal set of homogeneous generators called fundamental invariants.[20]

The fundamental invariants can be calculated with King's algorithm[22] implemented in the computer algebra system called Singular.[23] The Singular examples of A_2B and A_2B_2 molecular systems can be found in supplementary material. The fundamental invariants of molecules containing identical atoms including A_2B, A_3, A_2BC, A_2B_2, A_3B, A_4, A_2BCD, A_2B_2C, A_3BC, A_3B_2, A_4B, and A_5 molecules are calculated and the FIs of A_3B and A_4B are listed in supplementary material. The corresponding Fortran subroutines can be downloaded from https://github.com/kjshao/FI.

Now let us have a brief discussion on the relationship between FI, primary and secondary invariants, by using the following theorem given in Ref. 20, Chap. 3.

Let R^G denote the invariant ring of a finite group G. For any $f \in R^G$, there exists a set of invariants $P = \{p_i \in R^G, 1 \leqslant i \leqslant n\}$ (called as primary invariants), a set of invariants $S = \{\theta_j \in R^G, 1 \leqslant j \leqslant t\}$ (called as secondary invariants), and a unique set of polynomials $\eta_1, \eta_2, \ldots, \eta_t \in R[p_1, \ldots, p_n]$ such that

$$f = \sum_{i=1}^{t} \eta_i(p_1, p_2, \ldots, p_n) \cdot \theta_i. \tag{1}$$

Equation (1) is the expression of Hironaka decomposition,[20,24] which states that primary and secondary invariants together can be used as the generators of R^G. Here $R[p_1, \ldots, p_n]$ is the ring generated by primary invariants.

The number of primary invariants equals the number of internuclear distances n. Let $d_1, d_2, \ldots, d_n$ be the degrees of the primary invariants and $|G|$ be the order of the group G, then the number of secondary invariants $t = \prod_{i=1}^{n} d_i / |G|$.

In fact, secondary invariants can be divided into two parts, the irreducible secondary invariants and reducible secondary invariants. The reducible secondary invariants are the combinations of the irreducible secondary invariants, and most polynomials in the secondary invariants are reducible secondary invariants. Primary and irreducible secondary invariants together are adequate for generating R^G.

In the King's algorithm, the fundamental invariants can be obtained directly without the calculation of primary and irreducible secondary invariants. In other algorithms, the primary and irreducible secondary invariants should be calculated first and fundamental invariants are obtained further by removing some abundant primary invariants.[25,26] Table S1 of the supplementary material lists the number of primary, secondary, irreducible secondary, fundamental invariants, and invariant polynomials truncated by the highest degree of FIs supplementary material. As can be seen, the number of invariant polynomials increases very quickly with the degree of polynomial. In contrast, the number of FI first increases then decreases with the degree of polynomial, in a much mild way. As a result, the total number of FI is substantially smaller than that of invariant polynomials to the same degree, in particular for systems involving multiple identical atoms.

The Hironaka decomposition is used in the PIP method[3] of Bowman and co-workers because of its compact representation of PESs. PIP-NN[12,13] combines PIP and NN methods by using the degree truncated invariant polynomials as the input vector of NN. In PIP and PIP-NN methods, polynomials are truncated by a highest degree bound to reduce the calculation as the number of polynomials increases nonlinear with the increment of the degree bound. With a carefully selected degree bound which balances the number of PIPs and fitting performance, PESs with small root mean square error (RMSE) can be obtained.[3,27,28] However, as the basis is incomplete, some extra error may be introduced into the PESs. Actually, in order to reduce the number of polynomials as the inputs of neural networks, primary and secondary invariants can be adopted because they can be treated as generators and their number is far more less than all the invariant polynomials truncated with a highest degree bound.

The fundamental invariant neural network is a neural network with the FIs as its input vector. The fundamental invariants set has the minimal number of polynomials that generate the invariant ring R^G. Because neural network can approximate any function to arbitrary accuracy by increasing the number of neurons in hidden layers,[6] using the fundamental invariants as the input vector of the neural network, one can approximate any permutation invariant potential energy surfaces. The FI-NN method is mathematically precise and the network structure of FI-NN is flexible as it has less limitations in the input layer.

As seen from Table I, for the A_3B molecules, there are in total nine fundamental invariants with a maximum

TABLE I. Comparison of experimental and computed vibration energy levels (in cm^{-1}) obtained on PIP-IMLS PES and FI-NN PES. Values in sixth and seventh columns are the differences between the experimental and the PIP-IMSL, FI-NN values.

State	Sym.	Expt.	PIP.	FI.	δ_1	δ_2
(0000)	A_1		9708.56	9689.10		
(0001)	F_2	1310.76	1313.07	1310.18	−2.31	0.58
(0100)	E	1533.33	1534.51	1531.36	−1.18	1.97
(0002)	A_1	2587.04	2591.54	2588.04	−4.50	−1.00
(0002)	F_2	2614.26	2618.82	2614.85	−4.55	−0.59
(0002)	E	2624.61	2629.21	2623.09	−4.60	1.52
(0101)	F_2	2830.31	2833.70	2827.50	−3.39	2.81
(0101)	F_1	2846.07	2849.58	2843.31	−3.51	2.76
(1000)	A_1	2916.48	2917.88	2912.41	−1.40	4.07
(0010)	E	3019.49	3019.39	3014.73	0.10	4.76

degree of three. This number is substantially smaller than the number of PIP of 50 up to degree four used by Li *et al.* in constructing a PIP-NN PES for the OH$_3$ system,[13] also smaller than the number of PIP of 22 up to degree three used by Li *et al.* for the OH$_3^+$ PES[29] (one may note that these 22 PIPs include all the nine FIs given here). In this work, the nine FIs were further transformed to $y_i = f_i^{1/m}$ where m is the degree of the corresponding FI. The set of y_i was used as the input vector in the fitting procedure using neural network. We choose $x/\sqrt{1+x^2}$ as the activation function, because it is faster for evaluation than the popularly used hyperbolic tangent activation function, and a network with a structure 9-40-40-1 with 2081 weights.

The data set used in the fitting of FI-NN PES is the 16 814 points which were used in the construction of OH$_3$ PES (NN1 PES).[11] The energies were calculated using ROHF-UCCSD(T)-F12a/aug-cc-PVTZ. The overall RMSE of NN1 PES is 1.61 meV with a network structure 6-50-50-1 in the interaction region,[11] while the RMSE of the PIP-NN PES is 1.4 meV, and the maximum deviation is 36.3 meV with a network structure 50-20-80-1.[13] 90% of the points were used as the training set while the left 10% were used as the validating set. After multiple training, three PESs were selected with least RMSEs to be included in the final
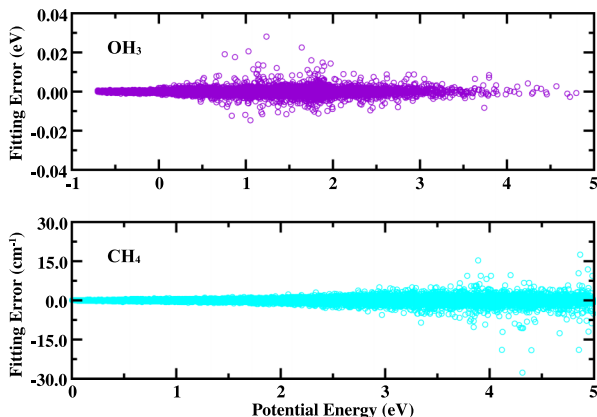
FI-NN PES with the averaging weights of 0.394, 0.332, 0.274. The three weights were randomly initialized and optimized with the descent steepest method to minimize the global RMSE. The training, validating, global RMSEs and the maximal deviation of the three PESs are (1.31, 2.09, 1.41, 29.62), (1.34, 2.26, 1.46, 27.05), (1.36, 2.46, 1.51, 29.55) in meV, respectively. The global RMSE, maximal deviation of the averaged FI-NN PES is 1.18 and 28.18 meV. The fitting errors are shown in the upper panel of Fig. 1.

On the new FI-NN PES, the reaction probabilities of the H$_2$ + O$_2$ and H + H$_2$O were calculated at $J = 0$ using full-dimensional quantum scattering. Figure 2(a) shows the comparison of the reaction probabilities for H$_2$ + OH → H + H$_2$O reaction on NN1 PES and FI-NN PES, while Figs. 2(b) and 2(c) are the comparisons of the reaction
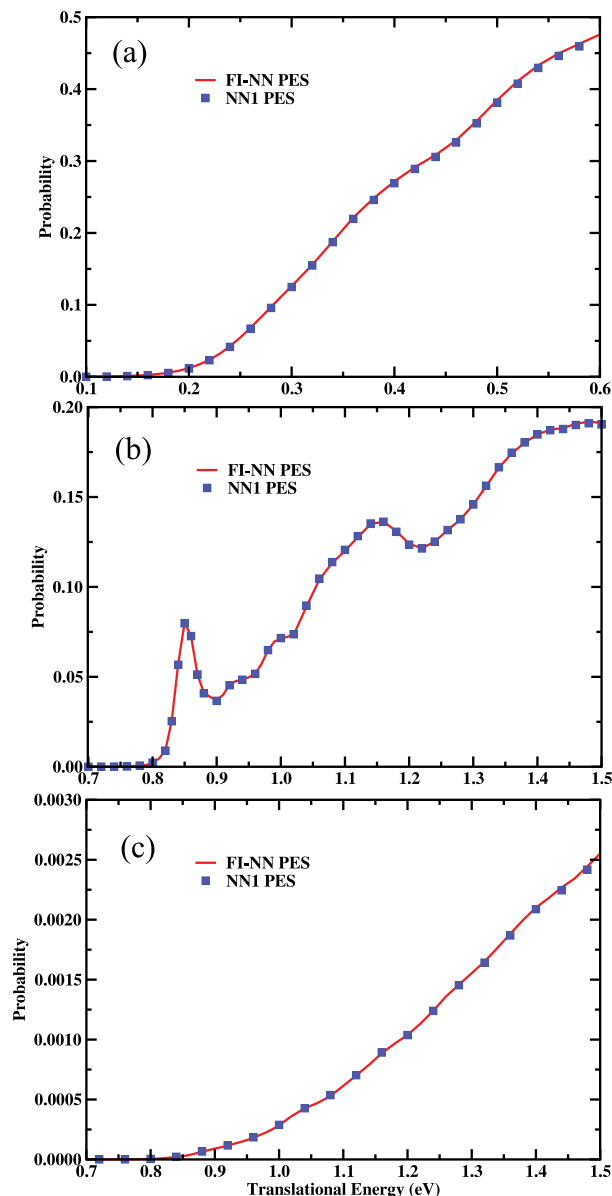
FIG. 2. (a) Comparison of the reaction probabilities for H$_2$+OH→H+H$_2$O reaction on NN1 PES and FI-NN PES. (b) Comparison of the reaction probabilities for H+H$_2$O→H$_2$+OH abstraction reaction on NN1 PES and FI-NN PES. (c) Comparison of the reaction probabilities for H+H′OH→H′+H$_2$O exchange reaction on NN1 PES and FI-NN PES.

FIG. 1. Fitting errors of the H$_3$O and CH$_4$ systems using FI-NN as a function of the corresponding potential energies.

probabilities for $H + H_2O \rightarrow H_2 + OH$ abstraction reaction and $H + H'OH \rightarrow H' + H_2O$ exchange reaction. Excellent agreement has been achieved for the above three reactions.

There are 31 fundamental invariants for the A_4B molecules with degree up to five. Similar to the fitting of OH_3 FI-NN PES, these FIs were transformed to $y_i = f_i^{1/m}$. The transformation of f_i to y_i can effectively accelerate the convergence both in OH_3 and CH_4 fitting procedure. The data set of the CH_4 molecule in this work contains 15 061 energy points which were calculated by ROHF-UCCSD(T)-F12a/aug-cc-PVTZ using MOLPRO 2012.1.[30] The maximal potential energy in the data set is up to 5 eV (40 328 cm^{-1}) relative to the CH_4 minimum.

A neural network structure 31-50-50-1 with 4201 coefficients and the $x/\sqrt{1 + x^2}$ activation function was used in the construction of FI-NN PES. Three best fitted PESs were included in the FI-NN PES and contribute to the final PES with the weights 0.421, 0.317, 0.262. The training, validating, global RMSEs and the maximal deviation of the five PESs are (0.92, 3.38, 1.38, 45.31), (1.06, 3.32, 1.46, 42.37), (1.03, 3.94, 1.59, 54.52) in cm^{-1}, respectively. The averaged FI-NN PES has a global RMSE, maximal deviation (1.04, 27.74) cm^{-1} with the energy up to 5 eV (40 328 cm^{-1}) and (0.52, 5.59) cm^{-1} with the energy up to 3.1 eV (25 000 cm^{-1}). The distribution of the fitting errors is shown in the lower panel of Fig. 1.

Majumder *et al.* constructed a PIP-NN PES of CH_4 in 2015 with 82 polynomials with a maximal degree of four.[31] Therefore, the FI generators in degree five were excluded in the polynomials used in that work, which may introduce some extra error into the PES due to the incomplete of basis. Another PES denoted as PIP-IMLS PES was also constructed and used as the benchmark PES in that work with the energies calculated at (AE)-MRCI-F12(Q)/CVQZ-F12 level. The RMSE of the PIP-IMLS PES is 1.01 cm^{-1} within 0-25 000 cm^{-1} and it is 2.14 cm^{-1} for the PIP-NN PES.

On FI-NN PES, the vibrational energy levels of CH_4 were obtained with a program for the 9-dimensional (9D) quantum mechanical calculation and compared with experimental and PIP-IMLS energy levels in Table I. For each degree of freedom, 8 potential optimized discrete variable representation (PODVR) grids were used for each degree of freedom. For more details, please see Ref. 32.

The deviations from the experimental energy levels of the calculated vibrational energy levels on FI-NN PES are smaller than corresponding ones on PIP-IMLS PES except the states (0100), (1000), and (0010). This difference may be caused by the different *ab initio* methods in the calculation of potential energies.

Fundamental invariants are the minimal generating set of an invariant polynomial ring. In this work, the permutation invariant of the PES was introduced to the neural network by using fundamental invariants as the input vector. FI-NN is mathematically precise and it can approximate any permutation invariant potential energy surfaces to arbitrary accuracy. The FIs of the molecular systems up to five atoms were generated by using Singular. The relationship between the fundamental invariants and other generators was discussed.

The potential energy surfaces of OH_3 and CH_4 were constructed with FI-NN with small RMSEs by using $x/\sqrt{1 + x^2}$ as the activation functions as it is much faster for evaluation than the hyperbolic tangent function. The accuracy of the PESs was confirmed with full-dimensional quantum dynamic scattering and bound state calculations.

Because of the reduction of the variables in the input layer, a FI-NN PES is faster for evaluation than the corresponding PIP-NN model PES as can be seen from the speed up shown in the last column in Table S1. In the comparison, the number of input FIs and PIPs for different systems is listed in Table S1. The weights in the neural network were randomly generated and the number of weights and neurons was selected to be nearly equal. There is no significant speed up for small systems, but for the A_3B_2, A_4B and A_5 systems, the speed up of FI-NN PESs is 1.87, 1.37, and 10.11, confirming the efficiency of FI-NN method. More details can be found in the supplementary material. For larger systems, we anticipate that the efficiency of FI-NN PES can be further improved.

See supplementary material for the examples of fundamental invariants and the speed up of FI-NN PES.

[1]G. C. Schatz, Rev. Mod. Phys. **61**, 669 (1989).

[2]T. Hollebeek, T. S. Ho, and H. Rabitz, Annu. Rev. Phys. Chem. **50**, 537 (1999).

[3]B. J. Braams and J. M. Bowman, Int. Rev. Phys. Chem. **28**, 577 (2009).

[4]Z. Xie and J. M. Bowman, J. Chem. Theory Comput. **6**, 26 (2010).

[5]J. M. Bowman, G. Czako, and B. N. Fu, Phys. Chem. Chem. Phys. **13**, 8094 (2011).

[6]K. Hornik, M. Stinchcombe, and H. White, Neural Networks **2**, 359 (1989).

[7]T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, J. Chem. Phys. **103**, 4129 (1995).

[8]D. F. R. Brown, M. N. Gibbs, and D. C. Clary, J. Chem. Phys. **105**, 7597 (1996).

[9]C. M. Handley and P. L. A. Popelier, J. Phys. Chem. A **114**, 3371 (2010).

[10]J. Behler, Phys. Chem. Chem. Phys. **13**, 17930 (2011).

[11]J. Chen, X. Xu, X. Xu, and D. H. Zhang, J. Chem. Phys. **138**, 154301 (2013).

[12]B. Jiang and H. Guo, J. Chem. Phys. **139**, 054112 (2013).

[13]J. Li, B. Jiang, and H. Guo, J. Chem. Phys. **139**, 204103 (2013).

[14]D. H. Zhang and H. Guo, Annu. Rev. Phys. Chem. **67**, 135 (2016).

[15]J. Chen, X. Xu, X. Xu, and D. H. Zhang, J. Chem. Phys. **138**, 221104 (2013).

[16]F. V. Prudente, P. H. Acioli, and N. J. J. Soares, J. Chem. Phys. **109**, 8801 (1998).

[17]H. Gassner, M. Probst, A. Lauenstein, and K. Hermansson, J. Phys. Chem. A **102**, 4596 (1998).

[18]J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401/1 (2007).

[19]J. Behler, Int. J. Quantum Chem. **115**, 1032 (2015).

[20]G. K. Harm Derksen, *Computational Invariant Theory*, Encyclopaedia of Mathematical Sciences, 2nd ed. (Springer, 2015).

[21]D. Opalka and W. Domcke, J. Chem. Phys. **138**, 224103 (2013).

[22]S. A. King, J. Symbolic Comput. **48**, 101 (2013).

[23]W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann, Singular 4–0–2—A computer algebra system for polynomial computations, 2015, http://www.singular.uni-kl.de.

[24]R. P. Stanley, Bull. Am. Math. Soc. **1**, 475 (1979).

[25]G. Kemper and A. Steel, "Some algorithms in invariant theory of finite groups," in *Computational Methods for Representations of Groups and Algebras* (Birkhauser Verlag Ag, Basel, 1999), Vol. 173, pp. 267–285.

[26]N. Borie, "Effective invariant theory of permutation groups using representation theory," in *Algebraic Informatics: Proceedings of the 6th International Conference, CAI 2015, Stuttgart, Germany, September 1-4, 2015*, edited by A. Maletti (Springer International Publishing, Switzerland, 2015), pp. 58–69.

[27]J. Li, J. Chen, Z. Zhao, D. Xie, D. H. Zhang, and H. Guo, J. Chem. Phys. **142**, 204302 (2015).

[28]J. Li and H. Guo, J. Chem. Phys. **143**, 221103 (2015).

[29]A. Li and H. Guo, J. Phys. Chem. A **118**, 11168 (2014).

[30]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.

[31]M. Majumder, S. E. Hegger, R. Dawes, S. Manzhos, X.-G. Wang, C. Tucker, J. Li, and H. Guo, Mol. Phys. **113**, 1823 (2015).

[32]Z. Zhao, J. Chen, Z. Zhang, D. H. Zhang, D. Lauvergnat, and F. Gatti, J. Chem. Phys. **144**, 204302 (2016).