1999-02-15

Determination of a methane intermolecular potential model for use in molecular simulations from ab initio calculations

Richard L. Rowley
rowley@byu.edu

Tapani A. Pakkanen

Determination of a methane intermolecular potential model for use in molecular simulations from *ab initio* calculations

Richard L. Rowley
Department of Chemical Engineering, Brigham Young University, Provo, Utah 84602

Tapani Pakkanen
Department of Chemistry, University of Joensuu, FIN-80101 Joensuu, Finland

The possibility of obtaining an accurate site-site potential model suitable for use in molecular dynamics (MD) simulations of methane from *ab initio* calculations has been explored. Counterpoise-corrected (CPC), supermolecule, *ab initio* energies at the MP2/6-311+G($2df$,$2pd$) level were computed for eleven relative orientations of two methane molecules as a function of C-C separation distance. C-C, C-H, and H-H interaction parameters in a pairwise-additive, site-site potential model for rigid methane molecules were regressed from the *ab initio* energies, and the resultant model accurately reproduced the *ab initio* energies. The model suggests that C-H attractions are dominant in weakly binding the methane dimer. CPC energies for methane trimers, tetramers, and a pentamer were also calculated at the same level. The results indicate that the n-mer energy per pair of interactions monotonically converges with increasing n, but that the assumption of pairwise additivity commonly used in MD simulations is reasonably valid. A limited number of higher-level calculations using MP4/6-311+G($2df$,$2pd$) and MP4/aug-cc-pVTZ were also performed to investigate the possibility of obtaining the intermolecular potential model from higher accuracy calculations without a substantial increase in computer resources. Results suggest that a Ne-methane probe method is not useful in this regard, but that limited, high-level computations, coupled with more extensive lower-level values, may be used to improve the model at minimal cost. © *1999 American Institute of Physics.* [S0021-9606(99)30207-5]

I. INTRODUCTION

The accuracy of thermophysical property values obtained from molecular dynamics (MD) simulations is primarily limited by the efficacy of the inter- and intramolecular potential used to model the fluid. It is therefore important to develop ways to obtain accurate intermolecular potentials. Recent successes in obtaining very accurate intermolecular potentials for dimers of inert gas molecules from post-Hartree-Fock *ab initio* calculations using large basis sets in conjunction with counterpoise corrections (CPC) for the basis set superposition error (BSSE) are most encouraging in this regard. Application of those techniques to more complex molecules, where relative orientation of the two molecules confuses the issue, requires additional study.

While it is probable that sometime in the future, *ab initio* calculations of intermolecular potentials ''on the fly'' during MD simulations[1,2] will become feasible, the CPU demand of such a method suggests that preassigned potential models will continue to play a significant role in MD simulations for many years to come. In this work, we develop analytical potential models for methane suitable for use in MD simulations from *ab initio* calculations. We also use the calculations to examine the efficacy of atomic site-site interactions and pairwise additivity, two common assumptions made in the analytical model potentials commonly used in MD simulations.

II. APPROACH

A. *Ab initio* calculation of intermolecular potential

Early attempts to derive intermolecular potential parameters from *ab initio* calculations focused necessarily on repulsions.[3,4] Repulsions can be calculated reliably even at the Hartree-Fock (HF) level. However, calculation of dispersion forces requires post-HF treatment of electron correlation. Woon and Dunning[5,6] found that the major limitation in describing weak interactions lies in the basis set rather than in the treatment of electron correlation, as long as Møller-Plesset (MP) perturbation theory is used with counterpoise-corrected (CPC) energies. Augmented correlation consistent basis sets (aug-cc-pVXZ) that contain a large number of polarization and diffuse functions were specifically developed to improve the quality of calculated dispersion energies.[6] Here, X refers to the level of complete shell functions added to the basis set. A quite definitive study on pairs of inert gas molecules using MP4/aug-cc-pVXZ showed that it is possible to obtain very accurate potential energies for atoms interacting through dispersion forces.[7] Interatomic potentials were calculated for He-He, Ne-Ne and Ar-Ar dimers using the supermolecule approach with CPC energies, and the results were in good agreement with the best known experimental values. The pair interactions were also studied as a function of MP perturbation order and basis set size. While MP4 calculations with very large basis sets (aug-cc-pVTZ or aug-cc-pVQZ where T=triple and Q=quadruple) were re-

quired for highly accurate potentials, calculations at MP2 with smaller basis sets also produced reasonably reliable results.

Two distinct approaches to obtaining pair energies have been attempted. The computationally less expensive method uses a test particle, generally an inert gas atom, to probe the intermolecular potential between a molecule and the probe. The interaction energy obtained by this probing is then fitted to an analytical pair potential. Generally the pair potential is modeled as a sum of all site-site potentials using the Lennard-Jones (LJ) equation for the site interactions. But, a combining rule, which relates the cross or heterogeneous interactions to the homogeneous interactions, must be assumed in order to extract the site-site interaction parameters for the molecular dimer from the calculated probe-molecule energies.[8–11] The more expensive method requires no such assumption as energies are directly calculated for a dimer of two identical molecules in different relative orientations, and site-site parameters are derived from the resulting dimer energy.[12–15]

Using the atomic probe method, Hill[11] determined that it was possible to derive reliable intermolecular potential model parameters based solely on *ab initio* calculations. He calculated MP2/aug-cc-pVTZ CPC energies for Ne-methane pairs in various relative geometries in order to regress parameters for a LJ-type analytical potential function. C-C and H-H interaction parameters were regressed from the Ne-C, Ne-H, and Ne-Ne CPC *ab initio* energies by using the LJ-parameter combining rules suggested by Waldman and Hagler.[16] The intermolecular potential parameters were consistent with empirically derived values and produced properties in MD simulations that were at least as good as the empirically derived models. Metzger, Gerguson, and Glauser[13] used the dimer approach with various semiempirical and molecular mechanics models in addition to MP2/6-311G($2d,2p$) *ab initio* calculations to determine methane-methane interactions for four different orientations. They found that the molecular mechanics methods were qualitatively consistent with the *ab initio* calculations, but that the semiempirical methods produced incorrect results in which the dimer was unbound. Previously, Novoa, Whangbo, and Williams[14] used MP2 and various smaller- to moderately-sized basis sets to determine the dissociation energy and equilibrium distance of the methane dimer. While prior HF calculations had predicted the dimer to be unbound, their MP2 results correctly showed the dimer to be bound with all the basis sets tested, and the dissociation energy of the dimer was in quantitative agreement with experimentally deduced values with MP2/6-311G($2d,2p$).

B. Standard MD intermolecular potential model assumptions

Current MD models are generally of the force-field variety with the potential represented as a sum of intra- and intermolecular potentials. Because of the time scale differences between translation and vibration modes, MD simulations often freeze internal modes, either assuming that the internal modes do not contribute to the desired property or preferring to calculate the contribution of the internal modes from statistical mechanics by assuming independence of vibrational and translational modes.

Two major assumptions are inherent in the model used in most MD simulations: pairwise additivity and the use of site-site interactions to model the full intermolecular potential energy.[17,18] Pairwise additivity assumes that the potential energy of molecule i is adequately approximated by a sum of isolated pair energies. Thus,

$$U_i = \sum_{j \neq i}^{N} U_{ij}, \tag{1}$$

where N is the number of molecules. This assumption is valid only if the pair interactions are independent of density; i.e., independent of the presence of other molecules in the fluid. As it is expected that the remaining $N-2$ molecules do affect the actual pair potential, we write

$$U_i = \sum_{j \neq i}^{N} [U_{ij} + \Delta U_{ij}(N)], \tag{2}$$

where $\Delta U_{ij}(N)$ is the change induced in the isolated pair potential, U_{ij}, by the presence of the other $N-2$ molecules.

The second, or site-site, assumption represents the distance, r, and angle, ω, dependence of the isolated pair potential as a sum of site-site potentials, each of which is spherically symmetrical. Thus, the isolated pair potential between molecules i and j can be represented by

$$U_{ij}(r,\omega) = \sum_{m=1}^{M} \sum_{n=1}^{N} U_{ij}^{mn}(r), \tag{3}$$

where U_{ij}^{mn} is the potential energy between site m on molecule i and site n on molecule j. Such potential models are particularly convenient for MD simulations because the angle dependence of the model is included implicitly in the various inter-site distances involved in the summation term of Eq. (3).

C. *Ab initio* potential model approach

In this work, a methane site-site potential model suitable for use in MD simulations of rigid methane molecules is developed. The CPC dimer-*ab initio* approach was used to calculate potential energy as a function of relative orientation and C-C distance between the two methane molecules. The calculated dimer energies were then used to develop an analytical potential model for methane consistent with the usual form of MD models as described above. In particular, site-site model parameters for C-C, C-H, and H-H interactions were determined directly from the potential energies obtained from the *ab initio* calculations. GAUSSIAN 94©[19] was used to perform all of the calculations.

The equilibrium geometry for a single, isolated methane molecule, constrained to Td symmetry and optimized with MP4/aug-cc-pVTZ, was used to model all methane molecules throughout this work. While geometry relaxation could be incorporated easily into the procedures used below, we are interested here in obtaining potential parameters for the rigid (fixed bond lengths and angles) methane model

TABLE I. Methane geometry from MP4/aug-cc-pVTZ.

Parameter	Definition	Value
$b_{\text{C-H}}$	C-H bond length	1.089 678 Å
$\angle_{\text{HCH}}$	HCH bond angle	109.4712°
d_{HCHH}	dihedral angles	120°; −120°

commonly used in MD simulations. Table I shows the methane geometry parameters obtained from the optimization.

Energy scans were performed on the methane-methane dimer using MP2/6-311+$(2df,2pd)$. CPC energies were obtained as a function of C-C distance, relative to infinite separation, for each of 11 different relative orientations of the two molecules. The relative orientations are depicted in Fig. 1. For ease of discussion, each orientation is given a two-letter acronym based on the approach of the two molecules, viewed as regular tetrahedrons. As depicted in the inset to Fig. 1, the notation used refers to the vertex (V), edge (E), and face (F) of a regular tetrahedron with the C nucleus at the center of the tetrahedron and a H atom at each vertex. Thus, the VV configuration is the head-on approach of two hydrogen atoms. Additionally, the rotation of one methane molecule relative to the other around the axis of approach can vary from staggered (St) to eclipsed (Ec) as shown in Fig. 1. We will refer to the six unique relative orientations of the tetrahedrons in Ec conformation as main orientations and to the five additional relative orientations studied here, involving rotation around the line of approach, as conformational variations of the main orientations. This approach is similar to that used previously,[13,14] but a larger basis set is used and a more complete set of orientations are computed.

III. RESULTS

A. Methane dimers

Potential energies obtained from the MP2/6-311+$(2df,2pd)$ calculations for the five main orientations are
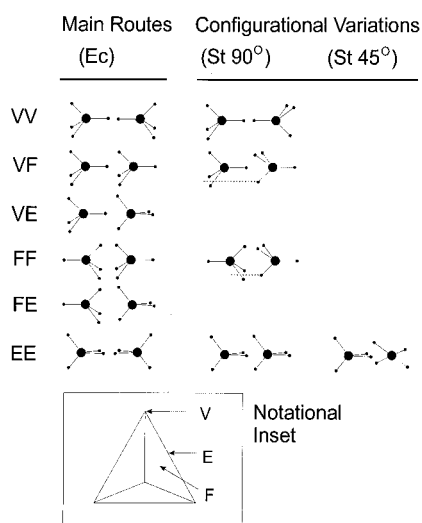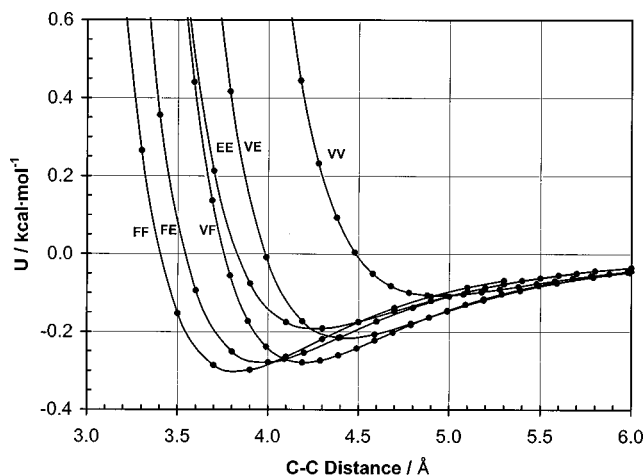

FIG. 2. Potential energy of methane dimer for the main orientations. Points are *ab initio* calculations; lines are smoothed curves.

shown in Fig. 2; the variation in energy due to configurational differences are shown in Fig. 3. Because of the inherent high degree of symmetry in the methane dimer, the relative geometries for which we have performed calculations cover a large portion of unique dimer geometry. Our results are consistent with the trends noted by Novoa, Whangbo, and Williams[14] in that the relative stability of the dimers is in the order $FF > FE \approx VF > VE > EE > VV$. Values of the energy minima are also in agreement with the orientations for which they made calculations with the 6-311G$(2d,2p)$ basis set. They report dissociation energies of 0.10 kcal/mol, 0.22 kcal/mol, and 0.30 kcal/mol for the VV(Ec), EE(St) and FF(St) orientations, respectively. These can be compared to 0.107 kcal/mol, 0.257 kcal/mol, and 0.334 kcal/mol for our 6-311+G$(2df,2pd)$ results for the same configurations. An experimentally deduced, isotropic potential[20] indicates that the effective methane-methane energy minimum is approximately 0.4 kcal/mol at a C-C distance of 3.85 Å. This dissociation energy is slightly larger than that obtained for even the most stable FF(St) configuration, but the reported dis-

FIG. 1. Relative orientations of the dimer pairs.
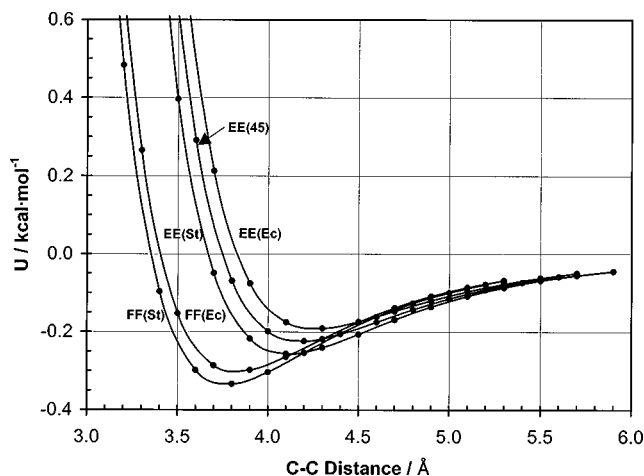
FIG. 3. Potential energy of configurational variations of main orientations. [VV(Ec)/VV(St) and VF(Ec)/VF(St) pairs are omitted because the configurational effect upon the energies is indistinguishable.] Points are *ab initio* calculations; lines are smoothed curves.

tance for the energy minimum is consistent with heavily *FF* weighted configurations. The relatively good agreement between the *ab initio* results and the experimentally deduced potential suggests that the analytical model developed from these results should provide quality MD simulation results.

B. Analytical site-site pair potential model

The reduction of the *ab initio* pair potentials into site-site potentials is not trivial. For any one orientation and C-C distance, there are 16 H-H, eight C-H, and one C-C interaction. Because bond distances are shorter than the distance over which dispersion interactions occur, none of these interactions dominate, and all 25 must be included in Eq. (3) in order to produce the overall potential described by the curves in Figs. 2 and 3. Thus, all potential parameters in the model must be regressed simultaneously.

Simultaneous regression of the C-C, C-H, and H-H potential parameters was attempted using various empirical, analytical potential models. Candidate models included LJ 6-12, LJ 6-9, LJ 6-8-12, exp-6, exp-6,8, and a modified Morse potential. The numbers in these models indicate the powers on the dimensionless reciprocal site-site separation distance. The modified Morse potential used was

$$U_{ij}^{mn}(r) = -\epsilon_{mn}(1 - \{1 - \exp[-A_{mn}(r - r_{mn}^*)]\}^2), \quad (4)$$

where ϵ_{mn}, A_{mn}, and r_{mn}^* are adjustable parameters. The LJ-*x* and exp-*x* models had difficulties in adequately modeling the wide orientational variation of the potential energy observed in Figs. 2 and 3. While all of the models can be used to fit the intermolecular potential of any one route, none of those used except Eq. (4) adequately described the energies of all eleven routes simultaneously within the framework of the site-site and pairwise additivity assumptions. To reduce the range of energies that these simple models are required to fit, large molecule-molecule repulsions that typically would not be observed in a MD simulation with adequately short time steps were not included in the data regression. This also helped prevent bias of the fitted equations toward the repulsions because of the dominant magnitudes of the repulsion potentials (hence their squared residuals in the regression) in this region. We chose to include all *ab initio* values below +3 kcal/mol in the regression. In all, 146 energies from all eleven orientations were included in the regression. The nonlinear, least-squares program utilized found a unique minimum with an average absolute residual of 0.08 kcal/mol. Comparison of the *ab initio* and fitted Morse-potential energies are shown in Figs. 4 and 5, and the best-fit parameters obtained for Eq. (4) are given in Table II.

The parameters in Table II should only be used as a complete set because there is still some coupling between the parameters. This is perhaps most easily seen from the estimation of the parameter standard deviations (obtained from numerical sensitivity coefficients) listed in parentheses below the value in Table II. Note that the fit is quite sensitive to the values of the C-C and H-H parameters, but is relatively insensitive to the C-C parameters. We assume that this is because the C-C interactions are essentially shielded in all orientations by the hydrogen atoms. The value of A_{CC} is still
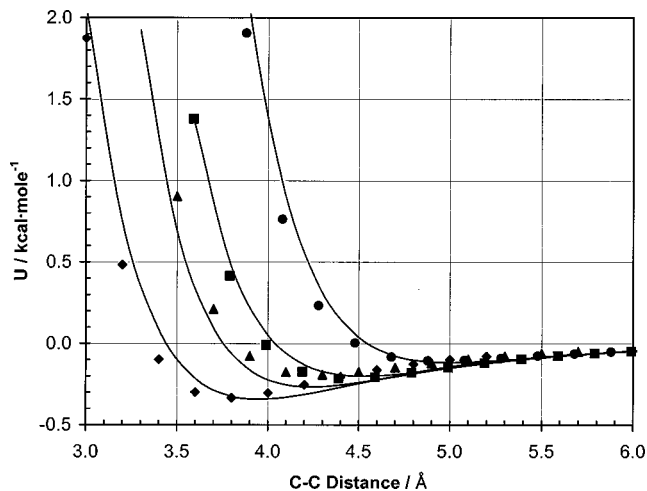

FIG. 4. Comparison of intermolecular potential for methane dimers from *ab initio* calculations (symbols) and from site-site model fitted with Eq. (4) (lines). Symbols: ●, *VV*; ■, *VE*; ◆, *FF*(St); ▲, *EE*(Ec). Note: The Ec and St variations of *VV* are indistinguishable.

quite well defined, but the other two parameters are coupled. In fact, r_{CC}^* could be fixed at a desired value and the other parameters in the model could be reregressed without a great loss in accuracy of the remaining fit. This is the meaning of the large estimated standard deviation associated with that parameter. It should also be noted, that this model is not intended for extrapolation to very short dimer separations for repulsions in excess of 100 kcal/mol, where the H-H repulsive potential in Eq. (4) can become unphysically attractive again. We chose to optimize the potential over the dimer separation ranges that would occur in normal MD simulations. Should shorter pair separations ever be required, additional *ab initio* energies should be included in the regression of a new model.

Figures 4 and 5 show that the overall fit with the simple, analytical, site-site, modified-Morse potential is reasonably good and could be used in MD simulations with rigid bonds
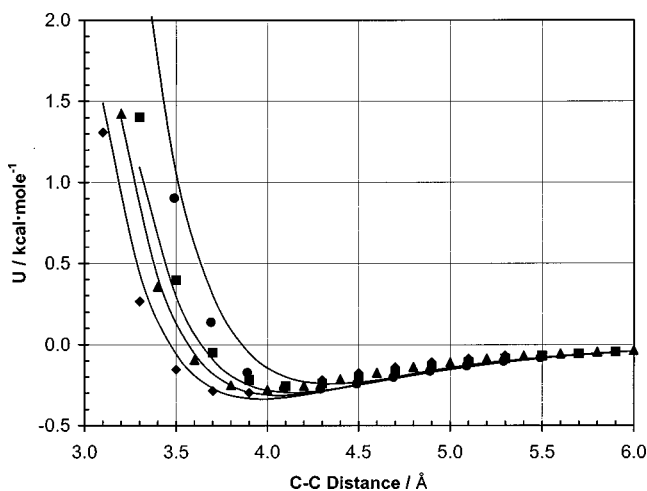

FIG. 5. Comparison of intermolecular potential for methane dimers from *ab initio* calculations (symbols) and from site-site model fitted with Eq. (4) (lines). Symbols: ●, *VF*; ■, *EE*(St); ◆, *FF*(Ec); ▲, *FE*. Note: The Ec and St variations of *VF* are indistinguishable.

TABLE II. Best-fit parameters for the modified morse potential, Eq. (4). Values in parentheses are estimated standard deviations of the parameters.

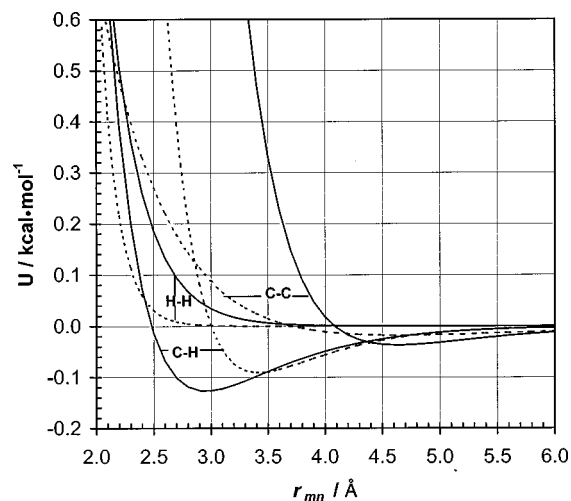| Parameter	C-C	C-H	H-H
(a) MP2/6-311+G($2df,2pd$)			
ϵ/kcal mol^{-1}	0.03714	0.12706	-5.93403
	(0.103)	(0.050)	(2.19)
A/Å^{-1}	1.25450	1.45976	3.39527
	(0.263)	(0.309)	(0.352)
r^*/Å	4.63470	2.95597	1.27724
	(16.8)	(0.333)	(0.157)
(b) MP4/aug-cc-pVTZ			
ϵ/kcal mol^{-1}	0.01793	0.09080	-18.936
	(0.006)	(0.005)	(3.83)
A/Å^{-1}	0.7544	1.6594	6.4206
	(0.058)	(0.031)	(0.328)
r^*/Å	4.63470	3.4220	1.3951
	⋯	(0.034)	(0.002)

FIG. 6. Site-site potentials obtained from dimer energies calculated using MP2/6-311+G($2df,2pd$) (solid lines) and MP4/aug-cc-pVTZ (dotted lines).

and angles, hopefully with improvement in multiproperty and thermodynamic state independence. However, some lose of fine-structure detail in the model can be observed in the transition region between repulsion and attraction. Though probably not warranted for calculations at this level of theory and basis set size, more complex potentials could be developed to capture this fine structure. For example, the dispersion coefficients C_6, defined by

$$U_{\text{disp},mn} = -\frac{C_6}{r_{mn}^6}, \qquad (5)$$

can be regressed for the C-C, C-H, and H-H interactions using only the clearly attractive potential values. In our case, all the *ab initio* values for which C-C distances are greater than 4.0 Å were fitted to Eq. (5). Whereas the ordering of the energies with respect to orientations was not completely matched by Eq. (4), regression of pure dispersion coefficients was found to correlate well the relative orientation and configuration dependence of the dispersion potential. Likewise, if only the data that are dominated by large repulsive terms are included in the regression, one obtains excellent fits of the repulsive potentials. We find that a simple exponential repulsive term

$$U_{\text{rep},mn} = B \exp(-D r_{mn}), \qquad (6)$$

where m and n can be C or H, adequately describes the strongly repulsive potential from 5 kcal/mol to several hundred kcal/mol. Parameters obtained for the dispersion and repulsive potentials are also shown in Table III. It is interesting to note that this analysis predicts that the C-C and H-H

TABLE III. Dispersion coefficients [Eq. (5)] and repulsive exponential parameters [Eq. (6)] regressed from appropriate portions of the MP2/6-311+G($2df,2pd$) data.

Parameter	C-C	C-H	H-H
C_6/kcal Å^6 mol^{-1}	-3959	815.9	-87.32
B/kcal mol^{-1}	12 505	2 742	849.2
D/Å^{-1}	-2.5829	-3.5955	-3.4782

interactions should be primarily repulsive in nature while the attractive potential is primarily supplied by the C-H dispersion.

A similar conclusion can be drawn from a plot of the site-site, modified Morse model regressed from the molecule-molecule energies. Figure 6 (solid lines) shows that the C-H interaction dominates the attraction between the molecules. It will be recalled that the molecule-molecule potential was fairly insensitive to the C-C interactions, and so there is less certainty about the attractive nature of those interactions. But, clearly the H-H interactions are nearly neutral in the normal dispersion region, and the C-H interactions are the dominant attractive potentials. This seems quite logical since both electrostatic field and Mulliken populations show that the C atom is significantly negatively charged and the H atoms are positively charged. Evidently, the Coulombic repulsion and dispersion attraction tend to cancel in the case of the hydrogen atoms as their interaction is fairly neutral. The potentially significant C-C dispersion attraction is also evidently moderated by the longer range Coulombic repulsion of the like-charged C atoms. But, the induced attractive charges and electron correlation combine in the case of C-H interactions, producing the dominant attraction between the two methane molecules. The dominance of this interaction is particularly evident when it is remembered that there are eight C-H interactions in the dimer and only one C-C. The 8:1 ratio of C-H to C-C interactions, coupled with the fact that the C-C distance is generally always larger than the dominant C-H or H-H separations, also helps explain the relative insensitivity of the molecule-molecule interactions to the C-C parameters.

C. Multibody interactions

The validity of the molecular pairwise additivity assumption, Eq. (1), can be directly tested with *ab initio* calculations of three or more methane molecules. As shown in Eq. (2), each actual pair potential can be thought of as the isolated pair potential plus a correction term for the effect of

TABLE IV. Relative orientations of clusters showing the type of cluster, the corresponding label in Fig. 7, the fixed base to which the Nth molecule is added, the orientation of molecule N, and the axis along which the distance of molecule N is varied.

Cluster	N	Fig. 7	Fixed base	Molec. N	Axis
Trimer1	3	A	FF(St)	F(St)	$-y$
Trimer2	3	B	FF(St)	V(St)	$-y$
Trimer3	3	C	FF(St)	F(Ec)	$+x$
Tetramer1	4	D	Trimer1	F(St)	$+y$
Tetramer2	4	E	Trimer1	F	$-z$
Pentamer	5	F	Tetramer1	F	$-z$

the other $N-2$ molecules upon that pair potential. The importance of the pair nonadditivity term was studied by performing MP2/6-311$+$G($2df,2pd$) energy calculations of N-mers or clusters of N molecules. In each case, $N-1$ methane molecules were fixed in a nearly optimal configuration and the energy of the N-molecule cluster was calculated as the Nth molecule was brought into the cluster along a path of fixed relative orientation. Table IV shows the five cluster orientations studied, each involving three to five methane molecules. Also shown in Table IV is the rigid, $N-1$, base cluster used for each energy scan. For trimers, the base was the FF(St) pair with the C-C distance fixed at 3.8 Å, corresponding to the FF(St) optimum pair geometry. The tetramer base, Trimer1, was the FF(St) base along the x-axis with the dimer center at the origin (marked with an X in Fig. 7) and a third methane in a F(St) orientation on the $-y$ axis
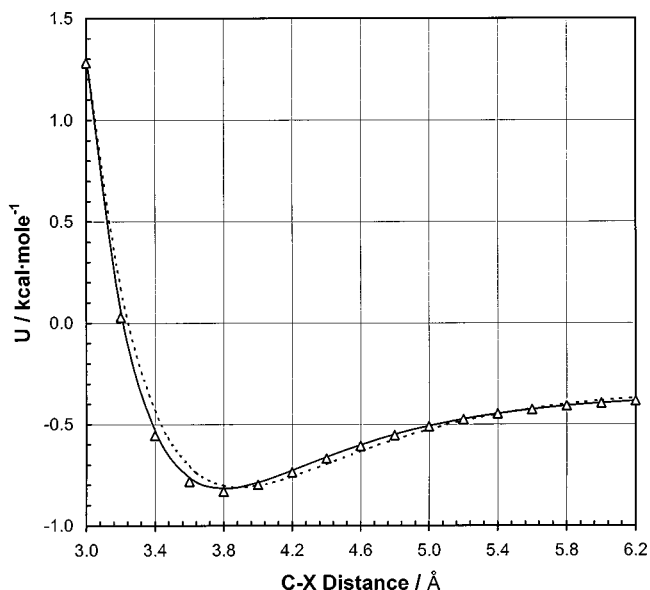
FIG. 8. Trimer1 CPC energy calculated from cluster ($\triangle$), pairwise additive potentials (—), and the pair site-site potential model given in Eq. (4) (---).

as shown in Fig. 7. The distance between the dimer center and the C atom of this third methane was also fixed at 3.8 Å, again corresponding to a minimum in energy for Trimer1. The pentamer base, Tetramer1, was Trimer1 with a methane in F(St) orientation added in the $+y$ direction, again with the distance between the C atom and the cluster center fixed at 3.8 Å, approximately corresponding to the minimum energy for Tetramer1.

An energy scan was performed as the Nth molecule approached the rigid base cluster along the orientations shown in Table IV and Fig. 7. The solid lines in Fig. 7 are intended to illustrate the base while the dotted line indicates the direction of the energy scan for the Nth molecule. Additionally, CPC energies for each binary pair of molecules in the cluster were computed. The BSSE for the clusters was corrected using a counterpoise correction for the N-mer. To perform this correction, the energy of each methane monomer, with the supermolecule basis set available to it, was calculated in the same configuration as for the supermolecule, but all methane molecules except the monomer were replaced with ''ghost'' methane molecules. In general, the CPC energy was found from

$$U_{N\text{-mer,CPC}} = U_{N\text{-mer}} - \sum_{i=1}^{N} U_i(M_i, G_{N-1}), \qquad (7)$$

where $U_i(M_i, G_{N-1})$ represents the energy calculated for a cluster of the same configuration as the N-mer, but with all molecules except i replaced with ''ghost'' methane molecules.

Results of the cluster calculations were compared with corresponding values for the sum of the pair potentials in the cluster. Figure 8 illustrates the agreement between the actual Trimer1 energies, pairwise additive pair potentials, and the pairwise additive site-site model developed in this work. The distances shown on the abscissa are between the C atom in the Nth methane molecule and the origin, or center of the
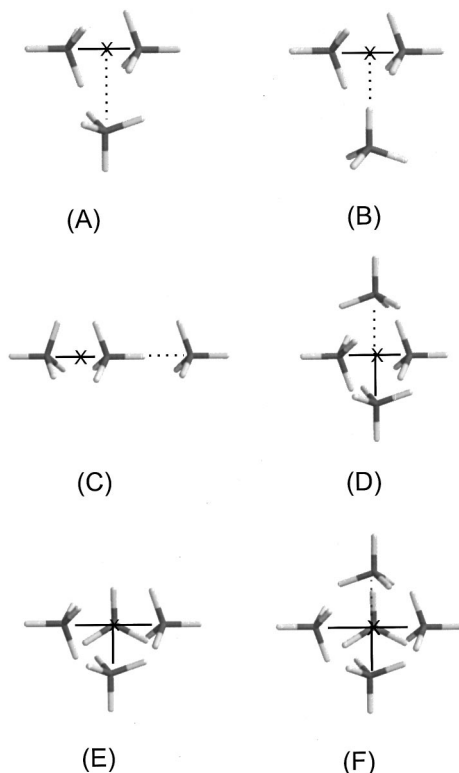

FIG. 7. Configurations of n-mer relative to origin (X) for (A) Trimer1, (B) Trimer2, (C) Trimer3, (D) Tetramer1, (E) Tetramer2, and (F) Pentamer. Solid line represents base cluster; dotted line shows the approach of the nth molecule.
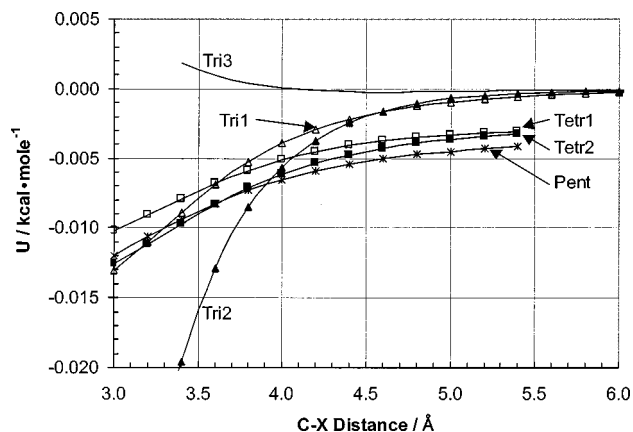
FIG. 9. Deviation of cluster potential from pairwise additivity per pair for Trimer1 ($-\triangle-$), Trimer2 ($-\blacktriangle-$), Trimer3 (—), Tetramer1 ($-\square-$), Tetramer2 ($-\blacksquare-$), and Pentamer ($-*-$).

TABLE V. CP-corrected pair energies obtained for Ne-methane and Ne-Ne dimers using MP4/aug-cc-pVTZ.

Orientation	$r_{\text{Ne-C}}/$ Å	$U(\text{MP2})/$ kcal mol^{-1}	$U(\text{MP3})/$ kcal mol^{-1}	$U(\text{MP4})/$ kcal mol^{-1}
Ne-V	3.206	1.2109	1.1604	1.0971
	3.736	0.0072	−0.0114	−0.0393
	4.265	−0.0645	−0.0699	−0.0818
	4.794	−0.0382	−0.0398	−0.0450
	5.323	−0.0199	−0.0204	−0.0229
	5.852	−0.0106	−0.0108	−0.0120
	6.381	−0.0060	−0.0060	−0.0067
Ne-F	2.646	1.7963	1.7525	1.6636
	2.911	0.4084	0.3756	0.3137
	3.175	−0.0143	−0.0309	−0.0733
	3.440	−0.1097	−0.1173	−0.1457
	3.704	−0.1077	−0.1109	−0.1299
	3.969	−0.0843	−0.0854	−0.0981
	4.233	−0.0613	−0.0616	−0.0701
Ne-E	2.800	1.8315	1.7601	1.6804
	3.320	0.0420	0.0195	−0.0182
	3.840	−0.0881	−0.0938	−0.1109
	4.360	−0.0551	−0.0564	−0.0641
	4.880	−0.0289	−0.0290	−0.0327
	5.400	−0.0154	−0.0154	−0.0172
Ne-Ne	2.400	1.1797	1.0817	1.0997
	2.660	0.2544	0.2111	0.2088
	2.920	0.0156	−0.0026	−0.0105
	3.180	−0.0334	−0.0410	−0.0488
	3.440	−0.0350	−0.0382	−0.0444
	3.700	−0.0272	−0.0286	−0.0332
	3.960	−0.0194	−0.0201	−0.0233
	4.220	−0.0136	−0.0139	−0.0162
	4.480	−0.0095	−0.0097	−0.0113
	4.740	−0.0067	−0.0068	−0.0080
	5.000	−0.0048	−0.0049	−0.0057

cluster. As can be seen, the energies of the cluster agree reasonably well with both the sum of the pair potentials and the site-site analytical model. As the clusters were generally built from a base cluster in which the molecules were located in potential well minima, the results would be expected to pertain to liquid-like densities. The results therefore suggest that nonpairwise additivity errors for liquid simulations are generally small, but not insignificant. In our case, it is evident from Fig. 8 that the nonadditivity error is smaller than the inaccuracy of the parameterized intermolecular potential, but it could become an important consideration as more accurate *ab initio* potential models are developed.

The difference between the cluster potential and the sum of the pair potentials per pair is shown in Fig. 9. The deviations from pairwise additivity indicate a distinct orientation dependence. Trimer1 shows larger deviations than Trimer2 in the repulsive region, but both are quite small in the attractive region. Trimer3, with molecules oriented in a line, is insignificant over the entire range. Pair nonadditivity is larger for the asymmetrical configuration of Tetramer2 than for the symmetrical orientation of Tetramer1, presumably because of partial cancellation of the multibody effect by the symmetry of the opposing molecules in the cluster. Interestingly, the results suggest that deviations per pair of interactions converge monotonically with increasing cluster size. While, it is not clear whether the second coordination shell of molecules would continue this trend or ameliorate the effects of the first shell, the results for Trimer3 suggest that the second shell may have little effect on the previous pairs.

D. Neon-methane probe

The molecule-molecule energy scans that we have used to obtain a model intermolecular potential suitable for use in MD simulations are expensive due to the large number of orientations possible, the large basis set requirement, the necessity of using post-HF methods for treating electron correlation, and the need to do BSEE corrections. While continued development of computational algorithms will help decrease the time requirements associated with individual energy calculations, the number of possible orientations and electrons associated with larger molecules currently pose

limitations to this approach for larger molecules. It is therefore worthwhile to further investigate the more inexpensive atom-probe approach for its accuracy in comparison to the dimer results reported here.

To this end, MP4/aug-cc-pVTZ calculations of Ne-methane dimer energies were calculated as a function of Ne-C distances, again along fixed relative orientations. The CPC energies calculated are given in Table V. The MP2 results were virtually identical to the results reported by Hill.[11] As can be seen from the MP3 and MP4 results, the higher level of theory tends to shift the potential to lower energies, making it more attractive in nature as electron correlations are handled better. Similar calculations were performed for the Ne-Ne dimer for data reduction purposes, and those results are also reported in Table V.

The trade-off for the time savings obtained by using the Ne probe approach is a lack of rigor in extracting the methane-methane potential from the Ne-methane energies. Following the procedure used by Hill,[11] the LJ type potential used in the CFF93 force field model,[20,21]

$$U_{ij}^{mn} = \epsilon_{mn}\left[2\left(\frac{r_{mn}^*}{r_{mn}}\right)^9 - 3\left(\frac{r_{mn}^*}{r_{mn}}\right)^6\right] = \frac{A_{mn}}{r_{mn}^9} - \frac{B_{mn}}{r_{mn}^6}, \qquad (8)$$

was used to regress the Ne-C and Ne-H interaction parameters from the data in Table V. The two forms shown in Eq. (8) are equivalent, and the relation between the two sets of

parameters is given by Hill.[11] In order to obtain C-C and H-H interactions from the probe results, a combining rule must be assumed. The combining rules proposed by Waldman and Hagler,[16]

$$r_{mm}^* = (2r_{mn}^{*6} - r_{nn}^{*6})^{1/6}, \quad \epsilon_{mm} = \frac{\epsilon_{mn}^2}{\epsilon_{nn}} \left(\frac{r_{mn}^{*2}}{r_{mm}^* r_{nn}^*} \right)^6, \quad (9)$$

were used by Hill to compute the C-C and H-H parameters from the Ne-C and Ne-H values, and a similar procedure was followed here. Hill found that initial attempts to fit the carbon and hydrogen parameters at the same time yielded parameters that were not physically reasonable. Regression of our data using the second form of Eq. (8) also gave a negative value for $B_{\text{Ne-H}}$, which yields unrealistic parameters in the first form. We believe the reason for the apparent problem is with the form of Eq. (8) and the methodology associated with the Ne probe approach. To avoid physically unrealistic model parameters, Hill assumed that energies for the Ne-V route were dominated by the Ne-H interaction and that those for the Ne-F route were dominated by the Ne-C interaction. This allowed the Ne-H parameters to be regressed independently from the Ne-V data and the Ne-C parameters to be regressed solely from the Ne-F route. However, this assumption is incompatible with the pair-wise additivity assumption used in MD simulations. This is particularly true for the Ne-F route in which the Ne-H interactions contribute as significantly to the potential as the Ne-C interaction. We therefore chose to force physically significant parameters by fixing one parameter and regressing the remaining three from the data. As the parameter r^* in Eq. (8) is the location of the potential minimum and physically represents an interaction size, we fixed the value of r^* for Ne-H at 3.000 Å to correspond with the approximate location of the minimum found for the Ne-V route. The geometry of this route places the Ne and H atoms head-to-head such that the diameters of the two atoms should dominate the location of the potential minimum.

An additional problem associated with the atomic probe method is the omission of the actual electrostatic distribution implicit in the dimer approach. As discussed above, the attractive nature of the C-H interactions and the neutral or repulsive nature of the H-H interactions found in this study are easily explained in terms of the calculated electrostatic charge distribution. Interactions between partially charged atoms and the neutral atomic probe do not adequately de-
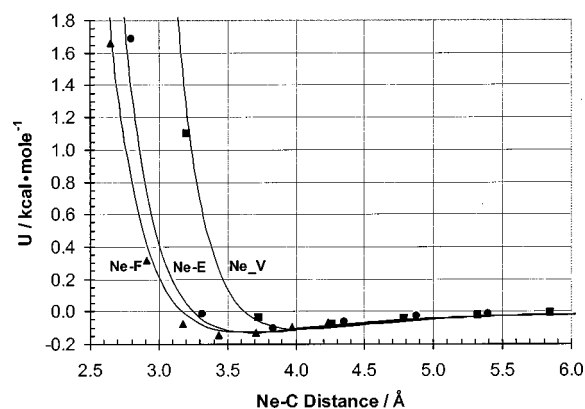

FIG. 10. Comparison of MP4 ab initio results (points) with fitted pairwise additive interatomic potentials (lines) for the Ne-V (■), Ne-E (●), and Ne-F (▲) orientations.

scribe the electrostatic atom-atom interactions in the dimer, nor can they be properly accounted for with an ad hoc combining rule. This is also the likely explanation as to why Eq. (8) was found to adequately fit the calculated Ne-CH_4 energies but not the calculated dimer energies.

The interatomic potentials obtained from the MP4 Ne-probe data are given in Table VI. Figure 10 shows that the three routes are fit well using the site-site model of Eq. (3). The interatomic pair potentials that result from this procedure are compared with the results obtained by Hill in Fig. 11. The C-C interaction is less attractive than Hill's MP2 results and the H-H potential is three times as deep, but of shorter range. The discrepancy illustrates the sensitivity of the resultant potentials to assumptions in the regression procedure. Additionally, both potential models indicate attractions between the hydrogen atoms, which our dimer results show to be incorrect. Based on ab initio dimer results, Novoa, Whangbo, and Williams[14] suggested that orientations with more H-H contacts were energetically favored. This statement is consistent with our dimer results that FF orientations are more energetically favored than VV, but this does

TABLE VI. Interatomic pair-potential parameters obtained from the MP4/aug-pVTZ results.

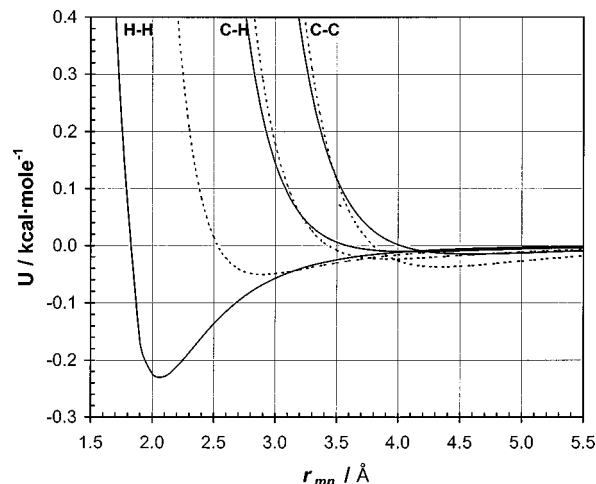| Pair	This work (MP4)		Hill[a] (MP2)	
	ϵ/kcal·mol^{-1}	r^*/Å	ϵ/kcal·mol^{-1}	r^*/Å
Ne-Ne	0.0634	3.337	0.0583	3.378
Ne-C	0.0205	4.192	0.0359	4.005
Ne-H	0.0541	3	0.0491	3.183
C-C	0.0149	4.5998	0.0374	4.3489
C-H	0.0105	4.1035	0.0237	3.9292
H-H	0.2303	2.0634	0.0506	2.8994

[a]Reference 11.

FIG. 11. Comparison of site-site potentials obtained from Ne-CH_4 probe calculations regressed as described in text (solid line) and from Hill (Ref. 11) (dotted lines).

not infer that H-H interactions are attractive. The same trends are explained by the strongly attractive C-H interactions found from the more rigorous methane dimer study. We believe that the reliance on ad hoc combining rules such as those in Eq. (9) constrains the regression such that the implied physics obtained for the site-site model are wrong. This statement is supported by the inability of the Ne probe method to obtain physically significant parameters when energies from all orientations were simultaneously regressed, and the disparity shown in Fig. 11. Although the probe method greatly reduces computer resource requirements, it is of questionable value.

IV. GENERALIZATION OF THE APPROACH

It seems clear from the dimer and Ne probe studies that potential energies for use in MD simulations can be calculated using dimers of the molecules. It is possible that small fragments of larger molecules can be used in the dimer calculations to provide information for the potential of the larger molecules. For example, the intermolecular potentials developed here may be appropriate for methyl groups in larger hydrocarbon molecules, and we are in the process of testing this idea. Even so, the required computer resources are large.

The methodology developed here is applicable to higher levels of theory and larger basis sets. Even more accurate potentials can be obtained by this method as computer speeds and quantum computational algorithms improve. Woon[7] found that MP4/aug-cc-pVTZ was required to yield high-accuracy potentials for inert gas dimers to within a few percent of experimental values. While we were able to carry out the Ne-methane study at this level, we do not have the resources at this time to develop as thoroughly as described above the methane pair potential with this large of a basis set. However, we were able to perform a limited study of the effects of MP level and basis set size upon the dimer potential. Calculations at five separation distances were repeated for each of the VV(Ec), FF(Ec), and EE(St) orientations using MP3/6-311+G($2df,2pd$), MP4/6-311+G($2df,2pd$), MP2/aug-cc-pVTZ, MP3/aug-cc-pVTZ, and MP4/aug-cc-pVTZ. The results are shown in comparison with the previous MP2/6-311+G($2df,2pd$) results in Fig. 12.

Both Fig. 12 and our results in the Ne probe study show a decrease in energy with increasing level of correlation treatment. Interestingly, this trend is sometimes monotonic with MP2>MP3>MP4, but more often follows the trend MP3>MP2>MP4. Woon[7] found similarly that dimer potential energies followed the former trend for He and Ne dimers, but the latter trend for Ar dimers. In all cases, the MP4 results were lower than MP2 for the same basis set. The difference between the MP2 and MP3 results is generally small. The difference between the MP2 and MP4 results is larger, mainly due to the triple calculations in the MP4 calculation, and fairly independent of the basis set.

Increased basis set size lowers the energy more than does increased level of correlation treatment. The most substantial decrease in energy is for the FF configuration (note the change in scale for the three figures). The minima in the MP4/aug-cc-pVTZ potentials are approximately −0.19,
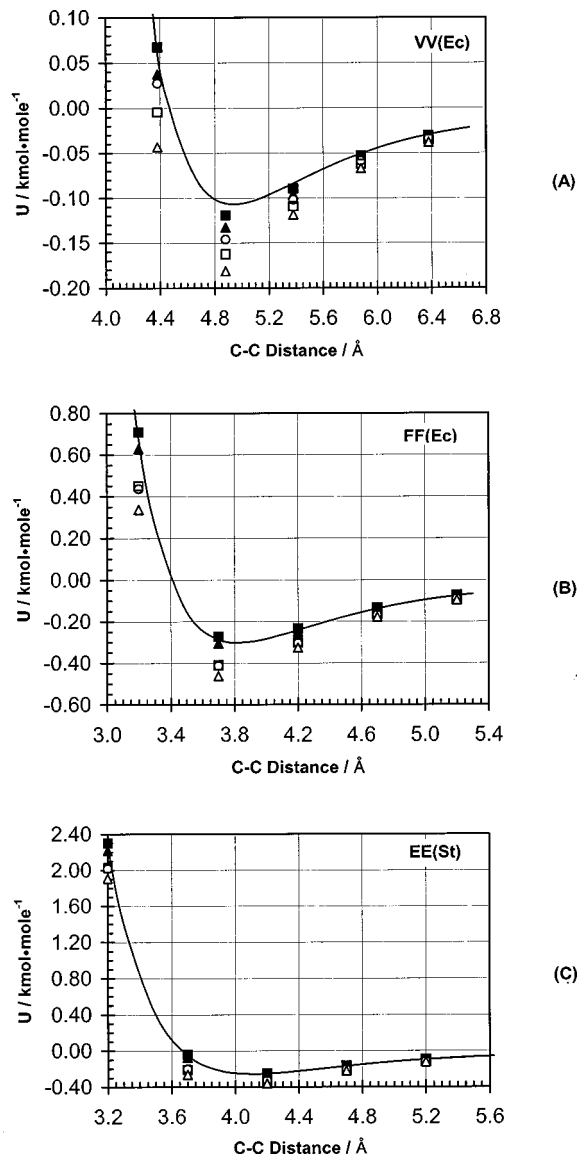

FIG. 12. Dimer energies for VV(Ec), FF(Ec), and EE(St) configurations for 6-311+G($2df,2pd$) (solid symbols and line) and aug-cc-pVTZ (open symbols) basis sets using MP2 (line, ○), MP3 (■,□), and MP4 (▲,△).

−0.40, and −0.47 kcal/mol for the VV, EE, and FF orientations, respectively. These values are in excellent agreement with the previously mentioned, experimental, isotropic potential minimum of −0.4 kcal/mol.

Often scaling is employed to correct for finite basis size or limited theory level. In this case the scaling factor would depend upon the configuration as well as the separation distance. Rather than attempt a scaling of the complete MP2/6-311+G($2df,2pd$) potential curves to the small number of MP4/aug-cc-pVTZ potentials calculated, we have simply done a reregression of the MP4/aug-cc-pVTZ data using the site-site model developed from the MP2/6-311+G($2df,2pd$) potentials. These do not constitute a large number of data points or orientations with which to do the regression, but the final MP4/aug-cc-pVTZ potential curves can be compared with the MP2/6-311+G($2df,2pd$) curves to visually ensure that the observed scaling effects occurred via the regression and that the relative energies of the various

orientations agree with the more complete MP2/6-311 +G($2df$,$2pd$) results. This is a way of using the completed smaller basis set work to guide the fit of a small number of more expensive calculations. The results of this procedure are shown in Fig. 6 (dotted lines) and the parameters obtained for the MP4/aug-cc-pVTZ potentials are given in Table II along with the parameters derived from the MP2/6-311+G($2df$,$2pd$) results. Because such a limited data set was available for regression, the r^* value for C-C (to which the data is fairly insensitive) was retained from the MP2/6-311+G($2df$,$2pd$) fit. The limited diversity of the configurations is also why the standard deviations computed by the regression program appear smaller than the previous fit. The essential features in Fig. 6 are similar for both potential models; the C-H interactions are attractive and the H-H interactions are neutral. Interestingly, the C-H attractive well is slightly shallower even though the overall MP4/aug-cc-pVTZ pair potential is more attractive. This results from the smaller distance over which H-H repulsions occur. As the regression of these MP4/aug-cc-pVTZ site potentials was performed from a very limited data set, one should avoid attaching additional physical significance to the shapes given. In particular, the large difference between the two C-C curves is due to the relative insensitivity of the C-C parameter to the data and the small number of points from different orientations that would tend to decouple the parameters. This suggests that caution should be used in regressing potential models from limited unique orientations.

V. CONCLUSIONS

MD simulations commonly use rigid, site-site, pairwise additive, intermolecular potential models for simulating fluid properties. We have examined here several aspects of the determination of such potential models from *ab initio* calculations. In particular, we have used MP2/6-311 +G($2df$,$2pd$) to calculate CPC energies of methane dimers in eleven different orientations as a function of separation distance. These energies were used to regress the parameters in an analytical site-site potential model that reproduces quite well the distance and orientation dependence of the *ab initio* energies.

The agreement of energies calculated from the site-site model with those for the variety of orientations obtained from the *ab initio* calculations indicates that the use of such models is appropriate and does not limit the accuracy of simulations. Calculations of methane trimers, tetramers, and a pentamer at the same level indicate that the pairwise additivity assumption is reasonably good at liquid densities, but illustrate that *N*-body effects calculated from clusters of molecules could be incorporated into more sophisticated models if needed. The model parameters obtained from this study indicate that the main attraction between methane molecules occurs as a result of the strong C-H interaction, and that H-H interactions are primarily repulsive. This has a logical explanation consistent with the electrostatic partial charges on those sites. However, these findings also suggest that a common practice in MD simulations of using ad hoc combining rules to obtain model parameters for unlike sites from known parameters for like sites is questionable. They also suggest that the use of such combining rules in the test-probe method makes that method considerably less accurate than the dimer approach.

We have also repeated a small number of calculations of CPC dimer energies using MP4/aug-cc-pVTZ for three of the eleven orientations. These were used to examine the dependence of the calculated energies on level of theory and basis set size. Basis set size was more significant than level of theory, but both tend to lower the energy. The minima in the potential energies obtained from these results were in excellent agreement with the experimentally determined, isotropic potential minimum for methane. These energies were also used to upgrade the potential model for the higher level of theory and larger basis set. We plan to use the newly developed analytical models in the future to extensively study the fluid properties of these models, as obtained from MD simulations, in comparison to all available methane data.

[1] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[2] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).

[3] U. Dinur and A. T. Hagler, J. Am. Chem. Soc. **111**, 5149 (1989).

[4] H. J. Böhm, R. Ahlrichs, P. Scharf, and H. Schiffer, J. Chem. Phys. **81**, 1389 (1984).

[5] D. E. Woon, Chem. Phys. Lett. **204**, 29 (1993).

[6] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).

[7] D. E. Woon, J. Chem. Phys. **100**, 2838 (1994).

[8] H.-J. Böhm, R. Ahlrichs, P. Scharf, and H. Schiffer, J. Chem. Phys. **81**, 1389 (1984).

[9] K. P. Sagarik and R. Ahlrichs, J. Chem. Phys. **86**, 5117 (1987).

[10] J.-R. Hill and J. Sauer, J. Phys. Chem. **99**, 9536 (1995).

[11] J.-R. Hill, J. Comput. Chem. **18**, 211 (1997).

[12] S. Tsuzuki, T. Uchimaru, K. Tanabe, and S. Kuwajima, J. Phys. Chem. **98**, 1830 (1994).

[13] T. G. Metzger, D. M. Gerguson, and W. A. Glauser, J. Comput. Chem. **18**, 70 (1997).

[14] J. J. Novoa, M.-H. Whangbo, and J. M. Williams, J. Chem. Phys. **94**, 4835 (1991).

[15] S. Tsuzuki, T. Uchimaru, and K. Tanabe, J. Mol. Struct.: THEOCHEM **307**, 107 (1994).

[16] M. Waldman and A. T. Hagler, J. Comput. Chem. **14**, 1077 (1993).

[17] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).

[18] R. L. Rowley, *Statistical Mechanics for Thermophysical Property Calculations* (Prentice–Hall, Englewood Cliffs, 1995).

[19] M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, GAUSSIAN 94, Revisions D.4 and E.2, Gaussian, Inc., Pittsburgh, Pennsylvania, 1995.

[20] R. Maple, M.-J. Hwang, T. P. Stockfisch, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler, J. Comput. Chem. **15**, 162 (1994).

[21] A. T. Hagler and C. S. Ewig, Comput. Phys. Commun. **84**, 131 (1994).