Pair approximation for polarization interaction and adiabatic nuclear and electronic sampling method for fluids with dipole polarizability

MILAN PŘEDOTA¹, PETER T. CUMMINGS²,³ and ARIEL A. CHIALVO¹,³

¹Department of Chemical Engineering, University of Tennessee, Knoxville, TN 37996-2200, USA
²Departments of Chemical Engineering, Chemistry, and Computer Science, University of Tennessee, Knoxville, TN 37996-2200, USA
³Chemical Sciences Division, High-temperature Aqueous Chemistry Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6110, USA

(Received 3 January 2002; accepted 15 February 2002)

The adiabatic nuclear and electronic sampling method (ANES), originally formulated as an efficient Monte Carlo algorithm for systems with fluctuating charges, is applied to the simulation of a polarizable water model with induced dipole moments. Structural, thermodynamic and dipolar properties obtained by ANES and a newer algorithm, the pair approximation for polarization interaction (PAPI), are compared with full iteration. With the best parameters, the inaccuracy of both approximate methods was found to be comparable with the uncertainty of the full iteration. The PAPI method with iteration radius equal to the second minimum of the oxygen–oxygen correlation function is, depending on the convergence tolerance, 10–15 times faster than the full iteration for 256 molecules, and yields very accurate structure and thermodynamics with deviation about 0.3%. When the iteration radius is increased to the cutoff distance, exact results are recovered at the cost of decreased efficiency. The ANES method with small nuclear displacements proved to inefficiently sample the configurational space. Simulations at low electronic temperatures with large nuclear displacements are inaccurate for up to 100 electronic moves, and increasing this number would make the simulations as slow as the full iteration. The most accurate and efficient adiabatic ANES simulations are those with infinite electronic temperature, large nuclear displacements and 1–10 electronic moves. The extra freedom of induced dipoles in the ANES method at high electronic temperatures modifies the observed dipolar properties; however, the question of whether the dielectric constant is also modified needs further consideration.

1. Introduction

In [1], we introduced the pair approximation for polarization interaction (PAPI), which proved to be a quite accurate and very fast method for Monte Carlo (MC) simulations of polarizable fluids with induced dipole moments. The full iteration procedure requires calculation of $O(N^2)$ interactions after each MC step, which makes simulation of polarizable models slower by $O(N)$ than simulations of their non-polarizable counterparts, where $N$ is the number of molecules. Our goal is to develop efficient approximate methods for MC simulations of polarizable fluids with induced dipole moments that would overcome the very long simulation times required by the exact methods. The PAPI approach scales as MC simulations of non-polarizable models, the pair correlation functions obtained by PAPI are indistinguishable from those obtained using the full iteration within the full scale graph, and the average configurational energy is shifted from the full iteration value by less than 0.4%, i.e. nearly within the standard deviations. In [1] we also investigated the single-update method following the approach of Medeiros and Costas [2], where only the induced dipole of a single molecule is modified during its displacement, keeping induced dipoles on all other molecules unchanged. We have proved that this method introduces strongly biased sampling of configurations, resulting in incorrect structure and thermodynamics, and should not be used [1]; this method was also questioned by Chen and Siepmann [3, 4].

In this paper we explore an alternative approximate method for the MC simulation of polarizable fluids with induced moments, which is a straightforward general-
ization of the adiabatic and nuclear electronic sampling (ANES) method, devised by Chen et al. [3] for models with fluctuating charges. Dipolar properties of the SCPDP model of water [5], including the dielectric constant (which has not been presented before), are calculated and the results of both PAPI and ANES simulations for a number of parameters are compared with those of full iteration, which serves as the exact method with which the approximate methods are compared.

2. Model

We investigated approximate methods for efficient MC simulations of molecular fluids with a single induced dipole moment per molecule, located typically in the centre of mass. In our simulations we used the SCPDP model devised by Chialvo and Cummings [5]. This model originates from the SPC and TIP4P non-polarizable models with fixed geometry and charges. However, in SCPDP the fixed partial charges are rescaled to give the correct dipole moment 1.85D of an isolated molecule and the Lennard-Jones (LJ) parameters are refitted to give the best agreement with experimental data. We used the set of parameters for oxygen–negative charge distance $|OM| = 0.2 \text{Å}$, which also has been studied by Jedlovszky and Richardi [6] using NVT-MC simulation with full iteration. We used $N = 256$ molecules and a cutoff radius of 9 Å. Long range interactions were accounted for by the LJ correction and reaction field technique with $\varepsilon_{RF} = 78$.

The polarizability is treated by a point–dipole interaction [5, 7]. There is a single induced moment $p_i$ located at the centre of mass of each molecule $i$, proportional to the total electric field $E_i$ at that point, that is a sum of the electric field from the charges on other molecules, $E_i^q$, and from the induced moments on other molecules, $E_i^p$:

$$
p_i = \alpha E_i = \alpha (E_i^q + E_i^p),$$

$$E_i^q = \sum_{\gamma=1}^{3} q_{\gamma}^j \frac{r_{i,j}}{|r_i - r_{\gamma}^j|},$$

$$E_i^p = \sum_{i\neq j}^{N} T_{ij} \cdot p_j,$$

$$T_{ij} = \frac{1}{r_{i,j}} \left( \frac{3 r_{i,j}^2}{r_{i,j}^3} - 1 \right).$$  

Equation (1) reduces to a simple expression:

$$U_{pol} = -\sum_{i=1}^{N} p_i \cdot E_i^q.$$  

The convergence criterion of the iteration procedure guarantees that the small deviation of induced dipoles from the values obeying equation (1) exactly introduces negligible difference in the polarization energy, and formulas (2) and (3) become equivalent, although each of them gives a slightly different result. Since equation (3) is easier and much faster to evaluate (does not require an extra loop over all pairs of molecules to calculate the energy after the convergence criterion was satisfied), it is used in the simulations for the calculation of the polarization energy [6, 9–15, 19] as opposed to the complete formula (2). However, it has to be borne in mind that
The electric fields and from neighbouring molecules, i.e. equation (2) is the generic formula valid for any induced moments \( p_i \), even in the cases when they do not obey the self-consistent equation. Since the derivative of equation (2)
\[
\frac{\partial U_{\text{pol}}}{\partial p_i} = -E_{\text{new}}^i - \sum_{j \neq i}^N (T_{ij} \cdot p_j) + \frac{p_i}{\alpha}
\]
is zero when the self-consistent equation (1) is satisfied, these values of induced dipoles minimize the polarization energy. For any other set of induced dipole moments, the polarization energy must be higher. This is not true when the simple formula (3) is used, because it is linear in induced dipoles and thus deviations from the self-consistent values of induced dipoles yield both negative and positive departure from the minimal polarization energy (cf. [1], figure 3). As long as the distribution of the deviations of induced dipoles from their self-consistent values is symmetric and narrow, this does not pose a problem and both the complete (equation (2)) and the simple (equation (3)) Hamiltonians give equal results.

3. Pair approximation for polarization interaction (PAPI)
The pair approximation for polarization interaction [1] updates the induced dipoles on the displaced molecule \( i \) and all molecules \( j \) within its cutoff sphere. The induced dipole \( p_j \) of the displaced molecule is calculated from the new electric fields \( E_{\text{new}}^i \) and \( E_j^p \), which are evaluated directly as a sum of contributions \( E_{i,j}^p \) and \( E_{i,j}^q \) from neighbouring molecules, i.e.
\[
E_{i,j}^{q,\text{new}} = \sum_{j \neq i} E_{i,j}^{q,\text{new}}; \quad E_{i,j}^{q,\text{new}} = \sum_{\gamma=1}^3 q_j^{\text{new}} \frac{r_{i,j}^{\text{new}}}{|p_i - r_{i,j}|^3},
\]
and
\[
E_j^{p,\text{new}} = \sum_{j \neq i} E_{i,j}^{p,\text{new}}; \quad E_{i,j}^{p,\text{new}} = T_{ij}^{\text{pol}} \cdot p_j^{\text{pol}}.
\]
The electric fields \( E_j^q \) and \( E_j^p \) on molecule \( j \) are modified by the displacement of molecule \( i \) and their new values are used to assign a new value of induced dipole \( p_j \). The electric field from charges \( E_j^q \) is updated using the old value \( E_{j,j}^{q,\text{old}} \), saved in an array, and the difference in the contribution \( E_{i,j}^{q,\text{new}} \) from molecule \( i \) to the total electric field from charges \( E_{i,j}^q = \sum_{j \neq i} E_{i,j}^{q,\text{new}} \) on molecule \( j \),
\[
E_j^{q,\text{new}} = E_j^{q,\text{old}} + \Delta E_j^q; \quad \Delta E_j^q = E_{i,j}^{q,\text{new}} - E_{i,j}^{q,\text{old}}.
\]
Only the interactions between molecules \( i \) and \( j \) need to be considered to update \( E_j^q \) accurately.
The change in electric field \( E_j^q \) from induced dipole moments is given not only by the change in \( p_i \), but also by the change in all other induced moments \( p_k \), because all induced moments respond to the displacement of molecule \( i \), i.e.
\[
E_j^{q,\text{new}} = E_j^{q,\text{old}} + \Delta E_j^q; \quad \Delta E_j^q = \sum_{k \neq j} E_{j,k}^{q,\text{new}} - E_{j,k}^{q,\text{old}}.
\]

Exact evaluation of this equation implies the calculation of all pair interactions (\( \propto N^2 \)), which is the most time consuming part of the full iterative procedure. PAPI takes advantage of the fact that most of the terms in the last sum are very small and can be neglected.

In the first iteration after a displacement of molecule \( i \), because all induced dipole moments still have their old values, \( \Delta E_j^p \) is given exclusively by the contribution from molecule \( i \),
\[
\Delta E_j^p = E_{j,i}^{p,\text{new}} - E_{j,i}^{p,\text{old}}.
\]

For the second and higher iterations, when the induced dipoles are already modified according to the updated electric fields in the previous iterations, equation (8) is no longer exact and we use instead an approximation of equation (7), which captures the essential contributions, but does not require calculation of all pair interactions. Since the strength of the polarization interaction decreases as \( r^{-3} \), the change in induced dipole moment \( \Delta p_k \) is small for molecules distant from the location of molecule \( i \). The term \( \Delta E_{j,k}^p \), proportional to \( \Delta p_k \), must also be small, especially when molecules \( k \) and \( j \) are far away, because \( \Delta E_{j,k}^p = T_{jk} \cdot \Delta p_k \) and \( T_{jk} \) is proportional to \( r_{jk}^{-3} \). Therefore, we consider for the update of induced dipoles only interactions between the displaced molecule \( i \) and all molecules within the cutoff radius of the new location of molecule \( i \) (as in the first iteration, see figure 1) plus all interactions between molecules \( j,j' \) which are distant from the old or new location of molecule \( i \) by less than a distance \( R_{\text{iter}} \), i.e.

![Scheme of the interactions used in PAPI](image)
\[ \Delta E'_j = \sum_{j' \neq i, j' \in R_{\text{iter}}}^{j} \left( E'_{j,j'}^{\text{new}} - E'_{j,j'}^{\text{old}} \right) + \sum_{j' \in R_{\text{iter}}} \left( E'_{j,j'}^{\text{new}} - E'_{j,j'}^{\text{old}} \right). \] (9)

The calculation of all interactions between molecules within distance \( R_{\text{iter}} \) guarantees proper handling of multibody interactions for short separations, which is essential for the correct update of induced dipole moments and the sampling of configurations. Because for a given \( R_{\text{iter}} \) the number of these interactions is limited and not a function of the number of molecules, PAPI scales linearly with the number of molecules \( N \); as for non-polarizable molecules. The \( N^2 \) dependence of the number of interactions evaluated by the exact procedure for polarizable interactions after each MC step is avoided. In [1] we found that the choice of \( R_{\text{iter}} \) equal to the location of the second minimum of oxygen–oxygen pair correlation yielded the best results in terms of accuracy and speed. For the SCPDP model under ambient conditions this choice corresponds to \( R_{\text{iter}} = 5.73 \text{ Å} \), i.e. on average there are 101 molecules in the cutoff sphere of radius 9 Å and 25 molecules within \( R_{\text{iter}} \).

From \( E'_{j}^{\text{new}} \) and \( E'_{j}^{\text{old}} \) the new induced moment \( p'_{j}^{\text{new}} \) is calculated according to equation (1). Since \( p'_{j}^{\text{old}} = \alpha(E'_{j}^{\text{old}} + E^q_{j}^{\text{old}}) \), there is no need to calculate or save the values of \( E'_{j}^{\text{old}} \); only the difference related to the move of molecule \( i \) is relevant, i.e.

\[ p'_{j}^{\text{new}} = p'_{j}^{\text{old}} + \alpha(E'_{j}^{\text{new}} - E'_{j}^{\text{old}} + \Delta E'_{j}). \] (10)

The update of both \( p'_{j} \) and all \( p'_{i} \)'s is iterated until convergence is achieved. The convergence criterion is the same as for the full iteration. The electric field from the charges, \( E'_{j} \) is updated throughout the simulation according to equation (6), because its value is needed to calculate the energy difference

\[ \Delta U_{\text{pol}} = -\frac{1}{2} \Delta \sum_{k=1}^{N} \left( p'_{k}^{\text{new}} \cdot E'_{k}^{\text{new}} - p'_{k}^{\text{old}} \cdot E'_{k}^{\text{old}} \right). \] (11)

used in the MC step. Here we use the ‘simple’ formula (3) for the calculation of polarizable energy, which is consistent with the use of this formula in simulations using the full iteration procedure, which we use as a benchmark. The complete formula (2) is inapplicable, because it is of order \( O(N^2) \). However, we have to be aware that in the case of PAPI the difference between the induced moments and their values obeying the self-consistent equation is larger than that when using the full iteration because of the approximation involved. The simple formula is used consistently throughout the whole simulation and thus it may be viewed as an effective or simplified Hamiltonian, which coincides with the complete one for the optimized induced dipoles.

The PAPI method can be combined with the full iteration. In [1] we found that even though the performance of PAPI is good if the system is evolved exclusively by the PAPI method, the accuracy can be improved without a significant increase in computational time when the induced dipoles are corrected by the full iteration after a specified number \( M \) of accepted MC steps. A choice of \( M \) from ranging between 10 and 100 was recommended.

In this paper, we present the results of four PAPI simulations. Three of them, denoted as PAPI_10, PAPI_100 and PAPI_∞ with the iteration radius \( R_{\text{iter}} = 5.73 \text{ Å} \) and \( M = 10, M = 100 \) and \( M = \infty \), respectively, where \( M = \infty \) denotes no full iteration.

In the fourth simulation, PAPI_9Å, \( R_{\text{iter}} \) was set equal to the cutoff radius of the potential, i.e. 9 Å, and no full iteration was performed \( (M = \infty) \); all dipolar interactions within the cutoff sphere of the displaced molecule were iterated to assign new induced dipoles on these molecules.

4. Adiabatic nuclear and electronic sampling method (ANES)

The ANES method has been described in detail by the authors Chen et al. [3, 16, 17] for the SPC-FQ fluctuating charge model of water. It stems from the idea of separation of nuclear and electronic degrees of freedom, the latter being subject in this method to a secondary temperature thermostat via a series of random changes of electronic variables, which follow each nuclear step, i.e. displacement of a molecule (either translation or rotation). This method is an improvement of the sequential fluctuating charge (SFC) algorithm [3, 4], in which either a nuclear or electronic move is performed as an MC step, and thus this algorithm might bias the sampling, favouring the initial state over the trial state in a nuclear move. The ANES method avoids this problem by simultaneous update of the electronic configuration during every nuclear move.

Though ANES can be used in the same way for both polarizable models with fluctuating charges or induced dipoles, for the second case it has not been studied so far. In our modification of ANES for models with induced dipole moments, the displacement of a molecule (nuclear step) and related change of induced dipoles (electronic step) are performed as follows.

- A randomly chosen molecule is displaced and the electric fields from charges and from induced dipoles are recalculated without any change of induced dipoles.

- Following the nuclear step, a series of \( R_{\text{elec}} \) electronic moves, random changes of induced dipoles on randomly selected molecules, is performed. The change of energy for each of these moves is calculated and the electronic move is accepted/rejected.
using a Boltzmann weight with an electronic temperature $T_{\text{elec}}$.

- The configuration resulting from the combination of nuclear and electronic moves is treated as the new trial configuration and is accepted/rejected using a Boltzmann weight with a temperature $T$ of the system.

Because of the nonzero electronic temperature, the electronic configuration of any molecule is not that of the ground state, but the distribution of electronic states should be given by the temperature $T_{\text{elec}}$, $1/T_{\text{elec}} = 1/T + 1/T_{\text{elec}}$, as a result of the double acceptance of electronic moves (at first as an independent electronic move and second during the acceptance of resulting trial configuration). Chen et al. proved theoretically that for models with quadratic coupling the sampling of the nuclear space is independent of $T_{\text{elec}}$ and identical to that of a Born–Oppenheimer (BO) system. For models with non-quadratic coupling, low electronic temperatures have to be used to keep the electronic state close to the ground state for each nuclear configuration. Chen et al. found for the SPC-FQ model [3], that when large trial displacements of a molecule are used at low electronic temperatures ($T_{\text{elec}} = 5\,\text{K}$, $0.5\,\text{K}$ and $0.05\,\text{K}$), large (up to 2000) numbers of electronic steps need to be carried out to guarantee proper sampling of electronic variables, and the method becomes as slow as the full iteration procedure. However, when the trial displacements were reduced from the usual values of 0.3 Å and 0.4 rad (giving an acceptance ratio of about 30%) to smaller values of 0.1 Å and 0.1 rad (acceptance ratio of about 67%), the number of electronic steps could be reduced to $R_{\text{elec}} = 10$. The explanation of this fact originates from the concept of the BO surface. For a large displacement, the random change in the induced dipoles is not efficient in jumping to the new BO skin of the new nuclear configuration. However, when the trial displacements are small, the BO skins of the new and old configuration are adjacent or even have partial overlaps and in such a case a lower number of electronic moves is sufficient for adequate sampling.

The most popular molecular models with fluctuating charges and the SCPDP model with induced dipoles have quadratic coupling, and therefore the distribution of electronic variables is independent of the nuclear variables and any electronic temperature can be used, including $T_{\text{elec}} = \infty$ [18]. In the latter case, all electronic moves are accepted and act merely as a special way of creating a new trial configuration, which is then accepted/rejected using the system temperature $T$. The extra electronic degrees of freedom contribute to the total energy; this additional energy must be subtracted from the configurational energy of ANES simulation to obtain the energy of the BO system of interest. For electronic variables independent of nuclear and other electronic variables the additional energy is given by the independent harmonic oscillator term $\Delta E = 3kT_{\text{elec}}/2$, coming from the three additional degrees of freedom of the induced dipole moment.

Unlike the full iteration or PAPI method, in which the induced dipoles are assigned according to the self-consistent equation (1), which guarantees that the assigned dipoles will be close to the values that minimize the polarization energy, in ANES the distribution of electronic states departs significantly from the minimized state of polarization energy, particularly for high $T_{\text{elec}}$.

Whereas in the full iteration and PAPI methods the energy is calculated after the new values of induced dipoles have converged, in ANES the energy change of each electronic move must be calculated. Obviously, in ANES the complete energy formula (2) must be used. The simple formula (3) without the self-consistent equation is unusable: any change in the induced moment of a molecule that has a positive scalar product with the electric field from charges on the same molecule makes the energy lower according to (3), which is unphysical. This formula can be used only to calculate the energy of optimized dipoles. Therefore, the complete formula for the polarizable energy must be used consistently in simulations using ANES. Because a single dipole is modified during an electronic move, ANES calculates $O(N)$ dipolar interactions during each electronic move.

5. Results

All the simulations were performed for the SCPDP model under ambient conditions, $T = 298\,\text{K}$ and $\rho = 1\,\text{g/cm}^3$ using $N = 256$ molecules. Extending our previous work [1], we explore the role of the tolerance of the convergence criterion in the full iteration and PAPI simulations; those with larger tolerance (lower accuracy) are identified by the symbol $\varepsilon$ in the tables and figures. We performed several ANES simulations for both large and small nuclear displacements with different electronic temperatures $T_{\text{elec}}$ and applying 1, 10 or 100 electronic moves. Simulations with large nuclear displacements, i.e. 0.3 Å and 0.4 rad, are distinguished in the tables and figures by the symbol $\Delta$ from those with small nuclear displacements (0.1 Å and 0.1 rad).

5.1. Full iteration and PAPI simulations

The full iteration is the simplest way of assigning induced dipole moments in simulations of polarizable fluids. Under the condition that the tolerance of the convergence criterion is small enough, this method should provide results indistinguishable from the exact ones, e.g. obtained by matrix inversion [3, 8], a method inapplicable to large scale simulations. For the discus-
sion of the approximate PAPI and ANES results, it is essential to obtain very accurate full iteration results to be able to reliably identify small differences among methods studied.

The convergence criterion for both full iteration and PAPI simulations in [1] was that the maximum change of any \( x \), \( y \) or \( z \) component of any induced dipole between subsequent iterations had to be smaller than 0.6% of the average value of the induced dipole. This tolerance is larger than that used by Jedlovsky [6, 10], namely ‘the iteration stopped when it changed the set of induced dipoles by less than 0.1%’, even if we assume that they considered the vector change of the induced dipole and not the change of each of its components. To resolve the role of the size of the tolerance in the convergence criterion, we performed full iteration and PAPI simulations both with the previously used large tolerance of 0.6% and with a new, smaller tolerance of 0.1%. The results obtained with the tolerance of 0.6% are distinguished from the more accurate ones by an extra symbol \( \varepsilon \) in the tables and graphs. The reduction in the tolerance resulted in an increase in the average number of iterations in each step from 3.2 to 4.4, and a corresponding increase in CPU time (see table 1). Because the first iteration of the full iteration method requires only the calculation of interactions of the displaced molecule (both at the old and trial positions) with all other molecules (equation (8)), making it much faster then the subsequent iterations in which dipolar interactions between all pairs must be calculated, the increase in the average number of iterations from 3.2 to 4.4 implies a CPU increase by a factor close to 3.4/2.2, i.e. the ratio of the number of slow iterations.

The detailed description of PAPI and the discussion of the results for different parameters was the subject of our previous paper [1]. However, because at that time we did not calculate the dielectric constant and other dipolar properties, for the purpose of this study we carried out four PAPI simulations for each tolerance. In two of the three simulations performed with \( R_{\text{iter}} = 5.73 \text{ Å} \), denoted PAPI_10 and PAPI_100, the full iteration of induced dipoles was applied after 10 and 100 accepted PAPI moves, respectively, while in the PAPI_\( \infty \) simulation no such update was applied. In simulation PAPI_9Å, all dipolar interactions between all molecules within the cutoff sphere of the displaced molecule are calculated, \( R_{\text{iter}} = R_{\text{cutoff}} = 9 \text{ Å} \), and no full update is performed.

In the full scale graph of the pair correlation functions, figure 2, no differences between the curves obtained by any full iteration or PAPI simulations can be observed. Therefore, we show in figure 3 only the area around the peak of the oxygen–oxygen correlation function, which is most sensitive to the differences. The reduction in the convergence tolerance of full iteration simulation has a small yet measurable effect: the shift of the peak by 0.01 Å to smaller separation and 0.03 higher. The performance of PAPI simulations relative to full iteration with the same tolerance follows the same pattern for both cases: the results for PAPI_10, PAPI_100 and PAPI_\( \infty \) are in very close agreement, with the latter two nearly identical, and all underestimate the height of the peak. PAPI_9Å gives excellent agreement with the full iteration, marginally overestimating the peak. The predictions of the average configuration energy \( E \), as seen in table 2, are analogous. The most accurate are PAPI_9Å simulations, which underestimate the energy by only 0.05 kJ mol\(^{-1}\), i.e. by an amount comparable with the statistical uncertainties;

Table 1. Comparison of CPU time in minutes per 1000 MC cycles for all the methods studied.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Tolerance</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.6%</td>
<td>0.1%</td>
<td></td>
</tr>
<tr>
<td>Full iteration</td>
<td>184</td>
<td>285</td>
<td></td>
</tr>
<tr>
<td>Full iter., full pot.</td>
<td>—</td>
<td>375</td>
<td></td>
</tr>
<tr>
<td>PAPI_10</td>
<td>20.3</td>
<td>24.4</td>
<td></td>
</tr>
<tr>
<td>PAPI_100</td>
<td>16.4</td>
<td>18.5</td>
<td></td>
</tr>
<tr>
<td>PAPI_( \infty )</td>
<td>15.9</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>PAPI_9Å</td>
<td>59</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>ANES, ( R_{\text{elec}} = 1 )</td>
<td>13.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANES, ( R_{\text{elec}} = 10 )</td>
<td>19.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANES, ( R_{\text{elec}} = 100 )</td>
<td>68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Pair correlation functions of SCPDP model at 298 K and 1 g cm\(^{-3}\). Within the scale of this graph, the lines for PAPI with \( R_{\text{iter}} = 5.73 \text{ Å} \) and no full updates (PAPI_\( \infty \)) and ANES with electronic temperature \( T_{\text{elec}} = 5 \text{ K} \), 10 electronic moves and small displacements (5 K, 10) are indistinguishable from the exact correlation functions obtained by full iteration. ANES simulation with large nuclear displacements is identified by \( \Delta \).
PAPI\textsubscript{10} overestimates the energy by less than 0.1 kJ mol\textsuperscript{-1} while PAPI\textsubscript{10} and PAPI\textsubscript{∞} overestimate the energy by slightly more than 0.1 kJ mol\textsuperscript{-1}.

The average values per molecule of the total dipole moment \( \mathbf{m} \), \( \mathbf{m} = \mu + \mathbf{p} \), induced dipole moment \( \mathbf{p} \), and cosine of angle between permanent and induced dipoles on each molecule, \( \cos(\mu, \mathbf{p}) \), shown in table 2, are in perfect agreement. To measure the dielectric behaviour of the whole system, we calculated the Kirkwood \( g_k \) factor [21–23],

\[
g_k = \frac{1}{N \langle m^2 \rangle - \langle M \rangle^2} = \frac{1}{N \langle m^2 \rangle^2} \left( \sum_{k=1}^{n} \frac{M_k^2}{n} - \left( \sum_{k=1}^{n} \frac{M_k}{n} \right)^2 \right),
\]

where \( m_i \) is the total dipole of molecule \( i \), \( M = \sum_{i=1}^{N} m_i \) is the total dipole of the simulation box and \( n \) is the number of measurements of this quantity, and the static Kirkwood factor \( G_K \),

\[
G_K = \frac{1}{N} \left( \sum_{i=1}^{N} \left( \frac{m_i}{|m_i|} \right)^2 \right) = 1 + \frac{1}{N} \left( \sum_{i \neq j} \cos \phi_{ij} \right), \tag{12}
\]

where \( \phi_{ij} \) is the angle between total dipoles of molecules \( i \) and \( j \). In the thermodynamic limit when \( \sum_{k=1}^{n} M_k \to 0 \), both these factors become identical. However, the total dipole moment of the simulation box is not zero, resulting in huge deviations (in the range 0–6) of the static Kirkwood factor \( G_K \), which is measured directly from the instantaneous configuration. More accurate determination of the Kirkwood \( g_k \) factor is given by the fluctuations of the total dipole moment of the simulation box, which we therefore prefer for our discussion and determination of the dielectric constant \( \varepsilon_r \) (according to [24], equation 24.3D)

\[
\frac{\varepsilon_r - 1}{\varepsilon_r + 1} \left( 1 - \frac{\varepsilon_r - 1 - \varepsilon_{RF} - 1}{\varepsilon_r + 1 + \varepsilon_{RF} + 1} \right)^{-1} = \frac{(\langle M^2 \rangle)}{36\varepsilon_0 V kT}
\]

as

\[
\varepsilon_r = 1 + \left( \frac{3\varepsilon_0 kT}{\rho g_k \langle m^2 \rangle} \right) \left( \frac{1}{2\varepsilon_{RF} + 1} \right)^{-1},
\]

where \( \varepsilon_0 \) is the permittivity of vacuum, \( \varepsilon_{RF} = 78 \) is the permittivity of the reaction field in our simulations, and \( m \) is expressed in SI units (Cm). The agreement between \( G_K \) and \( g_k \) is very good for PAPI simulations with length up to 3 million MC cycles. Because of the much slower calculation, the length of full iteration simulations was about 1 million cycles. The values of the dielectric constant do not depend on the tolerance used. In spite of the enormous CPU time spent on the full iteration simulations, we cannot determine \( g_k \) more accurately than in the range 1.8–2.1, corresponding to a dielectric constant \( \varepsilon_r \) = 72–90. Values from PAPI\textsubscript{9Å} are in agreement with the full iteration values of \( g_k \), while other PAPI simulations tend to underestimate the Kirkwood factor by 0.4, i.e. the dielectric constant by about 20.

So far we have compared PAPI simulations with full iteration simulations, which are performed in the standard way, i.e. evaluating the energy using the simple formula (3). We are not aware of any study of fluids with induced dipoles, that uses other than this formula, because of its simplicity, fast evaluation and the fact that in the limit of zero tolerance (infinite number of iterations) it is equivalent to the complete formula (2). However, because we work with approximate methods which are very close to the exact method, and moreover each of them uses a different formula for evaluating the energy, PAPI uses a the simple formula (3) while ANES uses the complete one (2), we found it interesting to investigate the behaviour of full iteration using the full potential formula (2) as well. Therefore we performed simulation denoted as ‘Full iter., full pot.’ with the smaller convergence tolerance 0.1%, in which the energy was calculated using equation (2) after completed iteration following each MC step, i.e. applying an extra loop over all molecular pairs to evaluate the energy. Needless to say, this simulation is very inefficient in terms of CPU time, see table 1. Although the length of this simulation is shorter, 600,000 MC cycles, the results are well equilibrated, except for the dielectric constant which might be underestimated and needs a longer run for a more reliable determination. The use of the complete potential formula modifies the results. The average energy is lower by 0.1 kJ mol\textsuperscript{-1} than that for a common full iteration, and the peak of oxygen–oxygen correlation function is higher. Both these findings make the agreement between this simulation and those using PAPI\textsubscript{10}, PAPI\textsubscript{100}, and PAPI\textsubscript{∞} less satisfactory,
Table 2. Average values of configurational energy $E$, electronic energy with respect to the ground state as observed in ANES simulations, $\delta E$, and as predicted by harmonic oscillator ideal term $\delta E'$, total dipole moment $m$, induced dipole moment $p$, cosine of angle between permanent and induced dipoles on each molecule $\cos(\mu, \mathbf{p})$, Kirkwood $g_k$ factor, static Kirkwood factor $G_K$, and dielectric constant $\varepsilon_r$ calculated from $g_k$. The maximum change $|\Delta \mathbf{p}|$ of an induced dipole in ANES and the acceptance ratios of electronic steps, $\text{Acc}_{\text{elec}}$ and of combined nuclear and electronic moves, $\text{Acc}_{\text{nuc}}$, are given in the last columns. Full iteration and PAPI simulations denoted by $\varepsilon$ use tolerance 0.6% of induced moments as convergence criterion, while the unmarked full iteration and PAPI simulations use tolerance 0.1%.

| Simulation | $E$ (kJ/mol) | $\delta E$ (J/mol) | $\delta E'$ (J/mol) | $m$ (D) | $p$ (D) | $\cos(\mu, \mathbf{p})$ | $g_k$ | $G_K$ | $\varepsilon_r$ | $|\Delta \mathbf{p}|$ (D) | $\text{Acc}_{\text{elec}}$ (%) | $\text{Acc}_{\text{nuc}}$ (%) |
|------------|-------------|------------------|------------------|---------|--------|-----------------|-------|-------|--------------|----------------|------------------------|------------------------|
| Full iter. $\varepsilon$ | $-40.75 \pm 0.02$ | $-2.82 \pm 1.00$ | $0.935$ | $1.8$ | $2.1$ | $72$ | $-30$ | $-30$ |
| Full iter. | $-40.79 \pm 0.05$ | $-2.82 \pm 1.01$ | $0.936$ | $2.1$ | $2.3$ | $90$ | $-30$ | $-30$ |
| Full iter. full pot. | $-40.89 \pm 0.03$ | $-2.82 \pm 1.00$ | $0.936$ | $1.6$ | $1.9$ | $61$ | $-30$ | $-30$ |
| PAPI_{10} $M = 10$, $\varepsilon$ | $-40.67 \pm 0.02$ | $0.1$ | $-2.81 \pm 1.00$ | $0.935$ | $1.5$ | $1.4$ | $56$ | $-30$ | $-30$ |
| PAPI_{100} $M = 100$, $\varepsilon$ | $-40.59 \pm 0.05$ | $1.0$ | $-2.81 \pm 1.00$ | $0.934$ | $1.5$ | $1.3$ | $56$ | $-30$ | $-30$ |
| PAPI_{∞} $M = \infty$, $\varepsilon$ | $-40.64 \pm 0.03$ | $2.5$ | $-2.81 \pm 1.00$ | $0.934$ | $1.5$ | $1.3$ | $56$ | $-30$ | $-30$ |
| PAPI_{9Å} $M = \infty$, $\varepsilon$ | $-40.80 \pm 0.02$ | $0.5$ | $-2.82 \pm 1.01$ | $0.935$ | $1.9$ | $1.8$ | $78$ | $-30$ | $-30$ |
| PAPI_{10} $M = 10$ | $-40.71 \pm 0.02$ | $0.1$ | $-2.82 \pm 1.00$ | $0.935$ | $1.4$ | $1.4$ | $51$ | $-30$ | $-30$ |
| PAPI_{100} $M = 100$ | $-40.67 \pm 0.03$ | $0.8$ | $-2.81 \pm 1.00$ | $0.935$ | $1.5$ | $1.5$ | $56$ | $-30$ | $-30$ |
| PAPI_{∞} $M = \infty$ | $-40.67 \pm 0.03$ | $0.8$ | $-2.81 \pm 1.00$ | $0.935$ | $1.5$ | $1.7$ | $56$ | $-30$ | $-30$ |
| PAPI_{9Å} $M = \infty$ | $-40.84 \pm 0.03$ | $0.3$ | $-2.82 \pm 1.01$ | $0.936$ | $1.8$ | $1.8$ | $72$ | $-30$ | $-30$ |

For $T_{\text{elec}}$ values:
- $T_{\text{elec}} = 0.5 \text{ K}$, $R_{\text{elec}} = 1$
- $T_{\text{elec}} = 0.5 \text{ K}$, $R_{\text{elec}} = 10$
- $T_{\text{elec}} = 0.5 \text{ K}$, $R_{\text{elec}} = 10, \Delta$
- $T_{\text{elec}} = 0.5 \text{ K}$, $R_{\text{elec}} = 100$
- $T_{\text{elec}} = 0.5 \text{ K}$, $R_{\text{elec}} = 100, \Delta$
- $T_{\text{elec}} = 5 \text{ K}$, $R_{\text{elec}} = 1$
- $T_{\text{elec}} = 5 \text{ K}$, $R_{\text{elec}} = 10$
- $T_{\text{elec}} = 5 \text{ K}$, $R_{\text{elec}} = 10, \Delta$
- $T_{\text{elec}} = 5 \text{ K}$, $R_{\text{elec}} = 100$
- $T_{\text{elec}} = 5 \text{ K}$, $R_{\text{elec}} = 100, \Delta$
- $T_{\text{elec}} = 25 \text{ K}$, $R_{\text{elec}} = 1$
- $T_{\text{elec}} = 25 \text{ K}$, $R_{\text{elec}} = 10$
- $T_{\text{elec}} = 25 \text{ K}$, $R_{\text{elec}} = 10, \Delta$
- $T_{\text{elec}} = \infty$, $R_{\text{elec}} = 1$
- $T_{\text{elec}} = \infty$, $R_{\text{elec}} = 1, \Delta$
- $T_{\text{elec}} = \infty$, $R_{\text{elec}} = 10$
- $T_{\text{elec}} = \infty$, $R_{\text{elec}} = 10, \Delta$
while the agreement with PAPI_9Å is still very good, although now the differences are of opposite sign. This finding indicates that the deviations arising from the approximation involved in PAPI are comparable with the uncertainties in the full iteration result itself.

5.2. ANES simulations at low electronic temperature

In this section we report the results of our simulations with $T_{\text{elec}} = 0.5\,\text{K}$ ($T'_{\text{elec}} = 0.499\,\text{K}$) and $T_{\text{elec}} = 5\,\text{K}$ ($T'_{\text{elec}} = 4.92\,\text{K}$). These low electronic temperatures were studied in [3], where the choice of low electronic temperature was motivated by an effort to keep the electronic degrees of freedom as close as possible to the ground state. An important parameter affecting the efficiency of ANES simulations is the maximum trial change of the induced dipole of a single molecule $\Delta p$ during a single electronic move. For low electronic temperatures, we had to keep $\Delta p$ very small to yield about 40% acceptance of a single electronic move. The acceptance rate of the combined nuclear and electronic steps is about 66% for small nuclear displacements and about 27% for large nuclear displacements. The acceptance ratios and $|\Delta p|$ are given in table 2.

The pair correlation functions of the SCPDP model at 298 K and 1 g cm$^{-3}$ are shown in detail in figure 4. The performance of ANES depends greatly on the parameters chosen, and the deviations of the less accurate simulations are visible even on the full scale graph of pair correlation functions (figure 2). All ANES simulations with large nuclear displacements (i.e. those with a symbol $\Delta$ in the legend) are less accurate than the simulations with small nuclear displacements, and result in peaks that are too pronounced, particularly for $R_{\text{elec}} = 10$ and $T_{\text{elec}} = 0.5\,\text{K}$, because the lower the electronic temperature, the greater the number of electronic moves needed for adequate sampling. The accuracy of simulations increases with the number of electronic moves $R_{\text{elec}}$. A single electronic move, $R_{\text{elec}} = 1$, is insufficient even for small nuclear displacements. The most accurate simulations are those with small nuclear displacements and $R_{\text{elec}} = 100$; the accuracy of simulations with $R_{\text{elec}} = 10$ is slightly inferior. The pair correlation functions for $T_{\text{elec}} = 5\,\text{K}$ are somewhat closer to the exact ones than those for $T_{\text{elec}} = 0.5\,\text{K}$. In all cases ANES overpredicts the height of the maxima of the pair correlation functions.

The resulting average configurational energies $\langle E \rangle$ are given in table 2. For ANES simulations, the values $\langle E \rangle$ given in table 2 are the average potential energies from the simulations minus the energy $\delta E_{\text{rev}} = 3kT'_{\text{elec}}/2$ of the extra degrees of freedom at $T'_{\text{elec}}$; i.e. $\langle E \rangle$ is the expected average energy of a BO system. There is a very good agreement between the simulations with small nuclear displacements and the exact result represented by simulation with full iteration. With large nuclear displacements the average energy is underpredicted, behaviour consistent with that of the pair correlation functions (the more pronounced the peak of oxygen–oxygen pair correlation function, the lower average energy). This finding indicates that the bias of configurational sampling introduced by ANES slightly favours low energy states.

Detailed information about the electronic energy distribution can be assessed by measuring the difference between the instantaneous energy $E$ of the system and corresponding minimized energy $E_{\text{full}}$ of the same nuclear configuration, $\delta E = E - E_{\text{full}}$. To calculate this distribution, the full iteration was performed in intervals of 1000 accepted MC steps without any effect on the ANES simulation. If the number of electronic moves were sufficient to sample the electronic degrees at $T'_{\text{elec}}$, the distribution of $\delta E$ would be centred around the harmonic oscillator value $\delta E_{\text{rev}} = 3kT'_{\text{elec}}/2$, which is given in table 2 and also shown in figure 5 as the vertical line for both temperatures. This is not the case for any ANES simulation with a low electronic temperature. The ANES simulations with $R_{\text{elec}} = 10$ are completely off the expected distribution and for $R_{\text{elec}} = 100$ the agreement is better; however, for $T_{\text{elec}} = 0.5\,\text{K}$ the distribution is still evidently off the expected average. Simulations with large trial nuclear displacements are completely offset towards higher energies. This indicates that the numbers of electronic moves used are insufficient to sample the system at the prescribed electronic temperature. Consequently, the distribution of electronic degrees of freedom might depend on the nuclear state of a molecule, which can result in biased sampling of both nuclear and electronic variables. In fact, if the true additional electronic energy, as quantified by $\delta E$ instead of the $3kT'_{\text{elec}}/2$ term, were subtracted from the average energy measured in ANES simulation, the

![Figure 4](https://example.com/figure4.png)

Figure 4. Oxygen–oxygen pair correlation function around its peak. ANES simulations are identified by the electronic temperature $T_{\text{elec}}$, number of electronic moves $R_{\text{elec}}$, and $\Delta$ for simulations with large nuclear displacement.
resulting total average energies would be even lower than the numbers given in Table 2. To obtain an estimate of how many electronic moves are required to reach the electronic equilibrium, we have calculated the evolution of the total energy as a function of the number of electronic moves, in the same way as done by Chen et al. (cf. [3], Figure 12). The electronic moves performed to show this evolution had no effect on the simulation. From Figure 6 we see that the energy decreases linearly up to about 1000 electronic moves, and not before 5000 moves does it reach the plateau, which corresponds to fluctuations of the electronic variables at $T_{0_{\text{elec}}}$. The energy of this plateau is shifted by the same amount as the distribution of $\delta E$ in Figure 5 is shifted from the energy corresponding to $T_{0_{\text{elec}}}$. Rigorously speaking, the distribution of electronic energy after an infinite number of electronic moves corresponds to the sampling temperature $T_{0_{\text{elec}}}$, while the distribution of states in simulation should follow $T_{0_{\text{elec}}}$, thanks to the acceptance of combined electronic and nuclear moves at temperature $T$. However, for these low electronic temperatures the difference between $T_{0_{\text{elec}}}$ and $T_{0_{\text{elec}}}$ is negligible. Both Figures 5 and 6 indicate that performing 10 or 100 electronic moves is definitely insufficient for the electronic sampling to relax after the nuclear move. Unlike in [3] for the SPC-FQ model, where this problem occurred only for large nuclear displacements, we observe this problem even for small trial displacements. Increasing the number of electronic moves is not a viable solution, because the CPU time increases linearly for large $R_{\text{elec}}$, with the result that the

![Figure 5. Distribution of the electronic energy $\delta E$ with respect to the ground state. The vertical lines denote the energy of harmonic oscillators at the electronic temperatures $T_{\text{elec}}$ corresponding to $T_{\text{elec}} = 0.5$ K and 5 K. The arrows help to identify simulation at $T_{\text{elec}} = 0.5$ K, with distribution falling incidentally into the region expected for simulations at $T_{\text{elec}} = 5$ K.](image)

![Figure 6. Evolution of energy $E$ with the number of electronic moves $N_{\text{elec}}$ for ANES simulations. Here $E$ is measured from the initial state.](image)

simulations with $R_{\text{elec}}$ about 400 would be as slow as the full iteration, see Table 1. Decreasing the nuclear displacements even further, which would reduce the number of electronic moves required for equilibration, is not a possible solution either, because as the nuclear displacements are reduced, a large number of MC steps is needed to sample the configurational space. In fact, even for the small displacements used here, i.e. 0.1 Å and 0.1 rad, the nuclear sampling is very inefficient and about 20 times more MC steps were needed to equilibrate the system and accumulate averages than for large nuclear displacements. This is the main drawback of the use of small nuclear displacements in ANES simulations.

All average dipolar properties are practically identical to those for the full iteration (Table 2). Only simulations at $T_{\text{elec}} = 0.5$ K with $R_{\text{elec}}$, too low (0.5 K, 1) and (0.5, 10, $\lambda$), result in slightly more parallel alignment of the permanent and induced dipoles. Even for our simulations of typical length of 1.4–3 million MC cycles, determination of $g_k$ is subject to errorbars of about 0.2. We observe a large scatter in the measured values of the Kirkwood factor from different ANES simulations, particularly for simulations with small nuclear displacements, which is a consequence of inefficient configuration space sampling. The values are most inconsistent for simulations at $T_{\text{elec}} = 0.5$ K, where we observe also large deviations between the values of Kirkwood factors $g_k$ and $G_K$ from one simulation. The inefficient sampling of the configuration space arising from the small nuclear displacements is critical for the determination of $g_k$, and limits the reliability of these predictions. Within these uncertainties and limitations, the dielectric constant of ANES simulations at low $T_{\text{elec}}$ is in accord with that of the full iteration simulation.
5.3. ANES simulations at high electronic temperature

To avoid the sampling deficiency at low electronic temperatures, Chen et al. [3] recognized that the sampling of nuclear coordinates for a system with quadratic coupling of nuclear and electronic variables should be independent of $T_{\text{elec}}$, including large temperatures and even $T_{\text{elec}} = \infty$.

We performed simulations with $T_{\text{elec}} = T = 298\, \text{K}$ ($T'_{\text{elec}} = 149\, \text{K}$), $T_{\text{elec}} = \infty$ ($T'_{\text{elec}} = T = 298\, \text{K}$) and one simulation at intermediate temperature $T_{\text{elec}} = 25\, \text{K}$ ($T'_{\text{elec}} = 23.1\, \text{K}$). At high temperatures and particularly for $T_{\text{elec}} = \infty$, $\Delta \mathbf{p}$ must be restrained to yield a reasonable acceptance ratio of the combined nuclear and electronic moves, which we kept at about 30% for small nuclear displacements and about 18% for large nuclear displacements.

Figure 7 shows that for simulations with large nuclear displacements, $T_{\text{elec}} = 25\, \text{K}$ and $T_{\text{elec}} = 298\, \text{K}$, the pair correlation functions reveal considerable inaccuracy and the sampling inefficiency of electronic degrees observed at low temperatures is still present. This is not surprising for $T_{\text{elec}} = 25\, \text{K}$, a temperature low enough to be considered in the previous part yet high enough to make the electronic energy non-negligible. As given in table 2, this energy is expected to be $288\, \text{J mol}^{-1}$ while from simulation it is $467\, \text{J mol}^{-1}$, because in the limited number of electronic moves the prescribed electronic temperature is not reached. In figure 8 we see that the distribution of electronic energy for $T_{\text{elec}} = 298\, \text{K}$ and $R_{\text{elec}} = 10$ is shifted towards high energies by 5% for large nuclear displacements, while for small nuclear displacements it is in accord with $T'_{\text{elec}}$. Pair correlation functions from all simulation at $T_{\text{elec}} = \infty$, regardless of the nuclear displacements and number of electronic moves, (which is 1 or 10), together with simulation $T_{\text{elec}} = 298\, \text{K}$ with small nuclear displacements and $R_{\text{elec}} = 10$, are in perfect accord with full iteration using the full potential and also very close to the common full iteration simulation using the simple potential formula.

Figure 7. As for figure 4 for ANES simulations with high electronic temperatures.

Figure 8. As for figure 5 for ANES simulations with high electronic temperatures.

The choice of $T_{\text{elec}} = \infty$ is a very special case, because all electronic moves are accepted and thus it simply reduces to a particular way of creating a new trial configuration, which is then accepted/rejected using the system temperature $T$ and the change of energy of the combined nuclear and electronic moves. In order to keep the acceptance ratio of a trial configuration reasonably high, the trial change $\Delta \mathbf{p}$ of an induced dipole on any molecule must be kept low, and even lower when $R_{\text{elec}}$ is increased. Otherwise, the sequence of completely random changes of induced dipoles would end up in configurations with unacceptably high energy. Note that it was the acceptance ratio of electronic moves, as opposed to the acceptance ratio of the combined nuclear and electronic steps, that limited $\Delta \mathbf{p}$ at low $T_{\text{elec}}$. Since for $T_{\text{elec}} = \infty$ all electronic moves are accepted, they cannot introduce bias for any $R_{\text{elec}}$ and nuclear displacements. This is confirmed in figure 8, where the distribution of electronic energies for all simulations at $T_{\text{elec}} = \infty$ fits a single curve given by thermal fluctuations at the system temperature $T = 298\, \text{K}$. Therefore, using large nuclear displacements and a single electronic move, $R_{\text{elec}} = 1$, seems to be the best choice based on sampling efficiency and CPU time. The average energies of all simulations with infinite electronic temperature lie between the values of the two full iteration simulations with smaller tolerance.

At this point we might conclude that ANES with high $T_{\text{elec}}$ samples the configuration phase space correctly. In fact, our findings so far plus the derivation in [3] prove it only for the nuclear part of the configurational space. The electronic sampling of ANES is different from that of the Born–Oppenheimer system, because the induced dipoles in ANES are free to fluctuate while there is no such freedom in the BO limit. This freedom of induced
dipoles increases with $T_{\text{elec}}$. As evident in table 2, even the average dipolar properties in ANES simulations differ from those of a BO system. The average induced dipole of a molecule is 8% higher than in a BO system, which is easy to understand as an effect of the extra electronic degrees of freedom, as shown in the appendix. For the same reason, the average cosine $\langle \cos(\mu, p) \rangle$ is smaller, i.e. the dominant parallel alignment of permanent and induced dipoles is less strict in ANES. Since these two effects have an opposite impact on the average total dipole moment $\langle m \rangle$, it is only by 1–2% higher in ANES than in the BO system. The extra degrees of freedom also seem to affect the Kirkwood factors, which generally are higher by up to 1.0 than for the BO system. Even though the measurement of $g_k$ is subject to large inaccuracy and very long runs are needed to calculate it, the consistency of $g_k$ values from the exact PAPI and ANES simulations at low electronic temperatures and for small nuclear displacements convinces us that the increase in $g_k$ and therefore in the dielectric constant in ANES simulations at infinite temperature is a measurable effect.

The idea of simulating a system that differs from that of direct interest is not uncommon, and is used successfully in a number of Monte Carlo methods, e.g. umbrella sampling or virtual insertions. The key to the success of such methods is a rigorous way of calculating the properties of the system of interest during the simulation of the different system. Our key question therefore must be: Can we correct the measured quantities in ANES simulations for its different electronic sampling with extra degrees of freedom and obtain the properties of a BO system? With regard to the potential energy, the correction based on the equipartition theorem accounts for the extra degrees of freedom very well. The correction suggested in the appendix might work quite well for the measured average dipoles and $\cos(\mu, p)$, in spite of its simplicity and the approximations involved; however, more sophisticated considerations are needed to elucidate this point and derive the relation between the values of dielectric constant measured in ANES and exact simulations.

6. Conclusions

We have presented a version of the adiabatic nuclear and electronic sampling (ANES) method applicable to polarizable fluids with induced dipoles, and extended our previous PAPI and full iteration simulations to the calculation of dielectric properties. Both PAPI and ANES proved to be successful approximate methods for faster simulation of these fluids. In all cases, the two methods yielded pair correlation functions that duplicated the shape of the correct ones and, with proper choice of parameters, were indistinguishable from the correct ones within the full scale of the graph. The average energy was always predicted within 1% and for the best choice of parameters, the approximation accounting for less than 0.3% difference in energy. Taking into account that the standard deviation of the measured energy is about 0.1%, the difference between the exact and approximate results is marginal. Because water is one of the most polarizable fluids, we must expect that the performance of approximate methods will be even better for fluids in which polarization effects are smaller. The greatest advantage of the approximate methods lies in simulations of large systems, because the exact MC methods for polarizable fluids are at least of the order $N^2$ interactions per single MC move, making these simulations very slow. The approximate methods scale as $O(N)$, i.e. as for MC simulation of non-polarizable fluids, and therefore their efficiency increases with the number of molecules.

In order to study small differences between the results of PAPI and ANES simulations with different parameters and for the determination of the dielectric constant, we carried out very long simulations, up to 3 million cycles for approximate methods and 1 million cycles for full iteration. Thanks to the high accuracy of our results, we could observe the dependence of the full iteration results on the convergence tolerance, which indicated that a tolerance of 0.6% of the average induced dipole might be too large. The reduction of this tolerance to 0.1% increased the CPU time for full iteration by 54%, but by only 8–20% for PAPI simulations with $R_{\text{iter}} = 5.73$ Å. We also investigated the effect of the complete and simple formulas for the evaluation of the energy in full iteration simulations. In spite of the fact that both these formulas become identical for induced dipoles that obey the self-consistent equation to which the iterated dipoles converge, each formula yielded slightly different results. This difference is negligible for most applications and the simple formula will definitely continue as the standard in simulations because of the significant time saving. However, we observe that full iteration results with the simple formula tend to be closer to PAPI results, which also uses the simple formula, while the ANES simulations are closer to the full iteration with the complete formula, which is used in ANES. In this light, PAPI and ANES with $T_{\text{elec}} = \infty$ are virtually as accurate as the full iteration itself, being referred to as the ‘exact’ method because the actual exact method, the matrix inversion, is prohibitively expensive for large systems. Another conclusion may be that an even smaller convergence criterion in the full iteration should be used if even higher accuracy is needed, of course at the cost of slower simulation due to more iterations.
There is no significant difference in accuracy or CPU time between PAPI simulations with \( R_{\text{iter}} = 5.73 \text{ Å} \) and full iteration update of all induced dipoles after \( M = 100 \) accepted MC steps (PAPI\_100) and simulations without any full iteration update (PAPI\_∞). The accuracy of PAPI with full iteration after \( M = 10 \) accepted MC steps is slightly higher in comparison with these two cases, as well as the CPU requirement, but it is questionable whether the combination of PAPI with full iteration updates and seeking the best frequency of updates \( M \) is worth the effort. Moreover, the detailed balance might be violated when full iteration is mixed with PAPI, although there is no positive evidence of this. Based on our latest results, we prefer PAPI simulations without any full iteration update, PAPI\_∞. The iteration radius \( R_{\text{iter}} \) affects both the accuracy and CPU time significantly. PAPI simulations with \( R_{\text{iter}} = 5.73 \text{ Å} \) are about 10–15 times faster for 256 molecules than the full iteration, depending on the convergence criterion, but overpredict the average energy by about 0.1 kJ mol\(^{-1}\). Simulations with iteration radius equal to cutoff radius, \( R_{\text{iter}} = 9 \text{ Å} \), are only 3.3 times faster for 256 molecules than the full iteration, but give results in perfect agreement with those from the full iteration. While the latter CPU saving might seem small, the ratio of CPU requirement becomes more favourable to PAPI\_9Å for polarizable systems with many molecules, for which this method might play a similar role to the most routinely used approximation in molecular simulations, the spherical cutoff of all fast decaying interactions. The only difference between PAPI\_9Å and full iteration is that PAPI\_9Å iterates all induced dipoles within the cutoff distance of only the displaced molecule, while full iteration does the same within a cutoff distance of each molecule. The determination of dielectric constant is consistent and the plateau of \( g_k \) is reached after about 200,000 cycles. Simulations with \( R_{\text{iter}} = 5.73 \text{ Å} \) predict too low a dielectric constant of about 56, while simulations with \( R_{\text{iter}} = 9 \text{ Å} \) give the correct value 72–78.

We have tested the performance of the ANES method for a range of electronic temperatures \( T_{\text{elec}} \) and frequencies of electronic moves \( R_{\text{elec}} \) and considered the effect of size of nuclear displacements. For low electronic temperatures, the use of large nuclear displacements leads to the largest departure from the exact results, even for \( R_{\text{elec}} = 100 \), and the increase of this number would slow down the simulation unacceptably. Simulations with small nuclear displacements are accurate and fast in terms of time per MC step; however, many more MC moves have to be carried out to sample the configurational space as efficiently as with large nuclear displacements. 10–20 times more MC cycles are needed for simulations with small nuclear displacements than for those with large ones, based on a comparison of time evolution of average energy and particularly of \( g_k \), since the point when it reaches its plateau seems to be a good estimate of the number of cycles needed to sample the configurational space sufficiently. This fact quite significantly deteriorates the actual efficiency of ANES simulations with small nuclear displacements, making them in fact as slow as the full iteration in terms of configurational sampling, and thus losing any advantage.

ANES simulations at infinite electronic temperature do not pose sampling problems, and the simulation with \( R_{\text{elec}} = 1 \) and large nuclear displacements was the fastest of all, giving the most accurate pair correlation functions among ANES simulations with an average energy within the uncertainties of the full iteration. If these properties are the only properties of interest, then this would be the best choice of ANES simulation. However, we have found that the extra freedom of induced dipoles in ANES modifies the dipolar properties of simulated system with respect to the Born–Oppenheimer system of interest. This effect is most evident for the average induced dipole per molecule and the average angle between permanent and induced dipole of a molecule. We have tried to explain this effect in the appendix using a simple assumption and ignoring most correlations. In doing so we have succeeded not only in linking the observed differences to the freedom of induced dipoles, but also even theoretically estimating these differences. However, our considerations are intended to be a simple explanation rather than a rigorous proof, i.e., a more sophisticated development is needed (if possible) to derive the correct formulas for the determination of the dipolar properties of a BO system from ANES simulations at infinite electronic temperature. Simulations at intermediate electronic temperatures combine the disadvantages of the two extreme cases (need for small nuclear displacements and modified dipolar properties) and therefore are not recommended.

Our considerations do not indicate that the dielectric constant measured in ANES should deviate from the correct one, though it is clear that, as stated in the appendix, our assumptions fail, particularly when cumulative effects like the total dipole moment of a simulation box are of interest. For example, the Kirkwood \( g_k \) factor from ANES simulations with \( T_{\text{elec}} = \infty \) seems to be too high, predicting a dielectric constant \( \varepsilon_r \) in the range 84–182. At the same time ANES simulation \((\infty, 10, \Delta)\) predicts a dielectric constant \( \varepsilon_r = 84 \) in accord with full iteration simulations. This indicates that the extra degrees of freedom in ANES at least introduce additional fluctuations to the determination of \( \varepsilon_r \) (if not an increase in the value), making the determination very difficult. With small nuclear displacements at low elec-
tronic temperatures, and therefore inefficient sampling of configurational space, we observe discrepancies between \( g_k \) and \( G_K \), particularly for \( T_{\text{elec}} = 0.5 \text{ K} \); however \( g_k \) is determined quite accurately except for the slow simulation (0.5 K, 100) for which the plateau of \( g_k \) was not reached within 1.4 million cycles with small nuclear displacements.

PTC’s and MP’s participation in this research was supported by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences, US Department of Energy. AAC’s participation was supported by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences, US Department of Energy, under contract DE-AC05-00OR22725 with Oak Ridge National Laboratory, managed and operated by UT-Battelle, LLC.

**Appendix**

The increase in average induced dipole moment and decrease in average cosine of the angle between permanent and induced dipole \( \cos(\mu, \mathbf{p}) \), due to the additional degrees of freedom of induced dipoles in ANES simulations, can be outlined using a simple argument as follows. Imagine that the induced dipole moment of a molecule \( i \) obeying the BO limit is \( \mathbf{p}^\text{BO} \) and the induced moment observed in ANES simulation is \( \mathbf{p}^\text{ANES} = \mathbf{p}^\text{BO} + \delta \mathbf{p} \), where \( \delta \mathbf{p} \) is the vector corresponding to the extra degree of freedom in ANES. Assuming that \( \delta \mathbf{p} \) has a random direction \( \gamma \equiv \angle(\mathbf{p}^\text{BO}, \delta \mathbf{p}) \), then for one specific size of this vector the average induced dipole moment observed in ANES simulation is given by integration over all angles \( \gamma \),

\[
\langle |\mathbf{p}^\text{ANES}^i| \rangle = \frac{1}{2} \int_0^\pi |\mathbf{p}^\text{BO} + \delta \mathbf{p}| \sin \gamma \, d\gamma \\
= \frac{1}{2} \int_0^\pi (|\mathbf{p}^\text{BO}|^2 + |\delta \mathbf{p}|^2 - 2|\mathbf{p}^\text{BO}||\delta \mathbf{p}||\cos \gamma |^{1/2}) \sin \gamma \, d\gamma \\
= |\mathbf{p}^\text{BO}| \left(1 + \frac{|\delta \mathbf{p}|^2}{3|\mathbf{p}^\text{BO}|^2}\right),
\]

and analogously,

\[
\langle |\mu^\text{ANES}^i| \rangle = \frac{1}{2} \int_0^\pi |\mu^\text{BO} + \delta \mu| \sin \gamma \, d\gamma \\
= |\mu^\text{BO}| \left(1 + \frac{|\delta \mathbf{p}|^2}{3|\mu^\text{BO}|^2}\right).
\]

The average cosine \( \cos(\mu, \mathbf{p}) \) equals (after some calculation)

\[
\langle \cos(\mu, \mathbf{p}^\text{ANES}^i) \rangle = \frac{1}{4\pi} \int_0^\pi \int_0^{2\pi} \frac{\mu^\text{BO}}{|\mu^\text{BO} + \delta \mu|} \sin \gamma \, d\gamma \, d\varphi \\
= \langle \cos(\mu, \mathbf{p}^\text{BO}) \rangle \left(1 - \frac{|\delta \mathbf{p}|^2}{3|\mathbf{p}^\text{BO}|^2}\right),
\]

This simple example shows that indeed the average induced dipole per molecule measured in ANES should be larger than that of the BO system while the opposite holds true for \( \cos(\mu, \mathbf{p}) \), and that the differences increase as the induced dipole moments are allowed to differ more from the values given by BO limit, i.e. with increasing electronic temperature. Substituting the values for ANES simulations at \( T_{\text{elec}} = \infty \), we obtain an estimate \( |\delta \mathbf{p}| = 0.46 \text{ D} \) and 0.45 D from the average dipole and cosine, respectively, while the less accurate estimate of \( |\delta \mathbf{p}| \) from the total dipole moment per molecule is about 0.5 D. The average deviation of induced dipoles in ANES from those of a BO system, \( |\delta \mathbf{p}| \), can be estimated theoretically from the classical harmonic oscillator formula [25]

\[
\langle |\mathbf{p}|^2 \rangle - \langle |\delta \mathbf{p}|^2 \rangle = 3\alpha kT_{\text{elec}},
\]

where \( \alpha \) is the scalar polarizability, yielding \( |\delta \mathbf{p}| = 0.42 \text{ D} \), since \( \langle |\delta \mathbf{p}| \rangle \) can be assumed zero. The good agreement between this prediction and our estimates from measured average quantities indicates that assuming isotropic distribution of \( \delta \mathbf{p} \) and ignoring any variability of \( \mathbf{p}^\text{BO} \) and \( \cos(\mu, \mathbf{p}) \) is acceptable for a rough estimate of the correction of the above mentioned properties per molecule. However, assuming independence of \( \delta \mathbf{p} \) on each molecule, i.e. no correlations between these vectors on different molecules as well as between these vectors and the orientations of the molecules, is definitely too crude an approximation to be able to assess the effect of ANES on the \( G_K \) factor. From equation (12), and assuming \( |m| = \text{const.} \), we would obtain

\[
G_K^\text{ANES} = \frac{1}{N} \left( \sum_{i=1}^{N} \frac{m_i^\text{ANES}}{|m_i^\text{ANES}|^2} \right)^2 \\
= \left( \frac{1}{N|m^\text{ANES}|^2} \left( \sum_{i=1}^{N} |m_i^\text{BO} + \delta \mathbf{p}| \right)^2 \right)^2, \\
\]

The second term is small (\( \approx 0.025 \)) and for a typical value of \( G_K \) it counterbalances the decrease in the first term due to the prefactor, resulting in effectively no difference between \( G_K^\text{ANES} \) and \( G_K^\text{BO} \). A more rigorous approach is definitely needed to elucidate this issue.
References


[18] Siepmann, I., private communication.


