



Entropy of Flexible Liquids from Hierarchical Force-Torque Covariance and Coordination

DOI:

[10.1080/00268976.2018.1459002](https://doi.org/10.1080/00268976.2018.1459002)

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Higham, J., Chou, S.-Y., Gräter, F., & Henchman, R. (2018). Entropy of Flexible Liquids from Hierarchical Force-Torque Covariance and Coordination. *MOLECULAR PHYSICS*. <https://doi.org/10.1080/00268976.2018.1459002>

Published in:

MOLECULAR PHYSICS

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.



Entropy of Flexible Liquids from Hierarchical Force-Torque Covariance and Coordination

Jonathan Higham^{a,b}, Szu-Yu Chou^{a,b}, Frauke Gräter^{c,d} and Richard H. Henchman^{a,b*}

^a Manchester Institute of Biotechnology, The University of Manchester, 131 Princess St, Manchester, M1 7DN, United Kingdom

^b School of Chemistry, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

^c Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany

^d Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Mathematikon, INF 205, 69120 Heidelberg, Germany

*email: henchman@manchester.ac.uk

Abstract

New theory is presented to calculate the entropy of a liquid of flexible molecules from a molecular dynamics simulation. Entropy is expressed in two terms: a vibrational term, representing the average number of configurations and momentum states in an energy well, and a topographical term, representing the effective number of energy wells. The vibrational term is derived in a hierarchical manner from two force-torque covariance matrices, one at the molecular level and one at the united-atom level. The topographical term comprises conformations and orientations, which are derived from the dihedral distributions and

coordination numbers, respectively. The method is tested on fourteen liquids, ranging from argon to cyclohexane. For most molecules our results lie within the experimental range, and are slightly higher than those by the 2PT method, the only other method currently capable of directly calculating entropy for such systems. As well as providing an efficient and practical way to calculate entropy, the theory serves to give a comprehensive characterization and quantification of molecular structure.

Introduction

Entropy is a fundamental quantity of a system that connects structure and thermodynamics. It is the logarithm of the weighted number of configurations and momentum states available to a system at a given set of conditions. Thus it indicates the extent of system flexibility and contributes to system stability. Despite its importance, entropy still cannot be calculated reliably. This is because the high dimensionality requires the evaluation of the relative probabilities of a massive number of configurations vastly greater than the number that can be sampled in a computer simulation. Calculating entropy is only straightforward when the probability distribution is simple, such as random, as in the ideal gas, or Gaussian, a good approximation for a crystal or a small, rigid molecule. However, most cases, including liquids and solutions, are intermediate between random and Gaussian and may possess any amount of complexity over many different length-scales. A common strategy is to calculate entropy differences from free-energy differences of one system relative to another,¹ matching the limitation of most experimental techniques which also can only provide entropy differences. While in principle exact, such approaches are limited not only by their technically and computationally demanding nature but forego valuable information on the state of interest not accessible by experiment.

The field of entropy calculations has progressed substantially over the past few decades.²⁻⁹ Probability distributions may be extracted from equilibrium simulations as a function of one or a few dimensions for a wide range of coordinates or order parameters. Such approaches, however, usually only give part of the total entropy, neglect correlations and require an

arbitrary choice of coordinates. Common coordinates measured include atom-atom distances and angles to construct the radial and angular distribution functions for liquids,^{10,11} or dihedral angles of flexible molecules.¹² Rather than using a simulation, entropy may be derived by integrating the Boltzmann factor over the coordinates of interest.¹³⁻¹⁶ Pairwise and higher correlations may be incorporated using the mutual information expansion in coordinates¹⁷ or conformation,¹⁸ with faster variants including only the largest correlation terms¹⁹ or correlations within a cutoff.²⁰ The k-nearest-neighbour methods address the difficulty of high-dimensional histograms by estimating the local density from the distance to the nearest configurations, either isotropically^{21,22} or anisotropically.²³

A long-standing and widespread approximation to account for linear, high-dimensional correlations is to assume a multi-dimensional Gaussian probability distribution for coordinates. The associated harmonic potential may be parametrized using either the Hessian at the energy minimum by normal mode analysis,²⁴ coordinate covariance by quasiharmonic analysis,²⁵ or forces.²⁶⁻³⁰ An alternative implementation of the harmonic approximation derives entropy from the vibrational density of states from the velocity autocorrelation function.^{31,32} It has been shown³³ that the matrix of Fourier transforms of the velocity autocorrelation function is equivalent to the force covariance matrix^{29,30} as well as the ensemble average of the Hessian. A notable advantage of harmonic approaches is that quantum effects can be incorporated easily,³⁴⁻³⁸ especially important for covalent bonds because of their high frequency.

A disadvantage of quasiharmonic analysis is that it averages over multiple wells³⁹ and ignores

non-linear correlations,^{40,41} a problem particularly acute for liquids,^{42,43} which requires special treatment to localise coordinate fluctuations.⁴⁴ Alternatives are to include the entropy from one-dimensional¹² or pairwise⁴⁵ probability distributions in place of the entropy from the diagonal self-covariance alone, so allowing for multiple minima. Other methods consider the entropy of multiple minima explicitly.^{13,18,20,46–56} Another approach, the 2PT method, includes a hard-sphere vibrational density of states for the zero-frequency vibration, with the full density of states from the velocity autocorrelation function^{32,57,58} to account for diffusive motion in liquids and give the correct gas-phase limit. A similar two-state strategy has been applied to account for the diffusive, anharmonic nature of low-frequency intramolecular vibrations.^{35,59} A variety of other strategies exist to calculate entropy, either from energy distributions,⁶⁰ transition probabilities to construct a molecule,⁶¹ or changes in low-frequency modes with respect to coordination number.⁶²

Notwithstanding all these developments, striking limitations remain. Most methods are aimed at single molecules or collections of small rigid molecules and are not suitable for systems of multiple flexible molecules. Only the 2PT method has yielded reasonable entropies for common organic liquids;^{58,63,64} the quasiharmonic method, when only accounting for intramolecular entropy, underestimated the entropy of liquid hydrocarbons.⁴⁰ Furthermore, many methods are limited to differences in entropy because they possess a bin-size dependence in the probability distributions, neglect quantum effects, or omit the integration over momenta, a non-trivial term if using internal coordinates.⁶⁵ Such a limitation has further amplified the common misconception that only entropy differences matter. Finally, despite

the multiscale nature of molecular structure almost all approaches consider a molecule's degrees of freedom at a single length scale. Only for normal mode analysis has grouping of atoms into rigid-body blocks been implemented.⁶⁶⁻⁷⁰

Here, we propose a method that can calculate the entropy of liquids of flexible molecules from the forces and coordinates generated in a molecular dynamics simulation. The philosophy of our approach is that it should be simple, fast, scalable, general, treats all molecules equivalently, gives a decomposition of entropy over all degrees of freedom of the system to explain the entropy obtained, and achieves comparable accuracy to that of the force-field used. The method builds on our earlier work for liquid water,^{27,37} organic liquids⁷¹ and single flexible molecules of united atoms.²⁹ Previously for water, vibrational frequencies had been extracted from force and torque magnitudes²⁷ or squared forces and squared torques³⁷ and the number of energy wells had been calculated using a generalized Pauling model^{52-55,72} originally developed for the residual entropy of ice.⁷³ For single, flexible, united-atom molecules, vibrational frequencies had been extracted from the eigenvalues of the force covariance matrix so as to account for correlations between covalently bonded atoms.²⁹ Here, for liquids of flexible molecules, we treat each molecule as a covalently bonded collection of rigid-body united atoms in the mean-field of its neighbouring molecules. We derive vibrational frequencies from the eigenvalues of a force-torque covariance matrix at the united-atom level and from another force-torque covariance matrix at the molecule level so as to capture flexibility at different length-scales. As well, we include terms to account for multiple orientations and conformations.

Theory

Entropy Decomposition Using a Hierarchy of Particle Sizes

The strategy here to calculate entropy is to partition configuration space into smaller, lower-dimensional regions which can be integrated over independently. In particular, we make the longstanding approximation of decomposing configuration space into energy wells of a given size and number, but we do so by treating the system as particles at a hierarchy of length scales. In high dimensions determining the weighted number of discrete energy wells is much more tractable than tackling continuous coordinates in the full configurational integral. Rather than concerning ourselves with the size of every distinct energy well, as would be done in normal mode analysis, we determine molecular vibrations averaged over all energy wells, similar to quasiharmonic analysis. This is helped by using the force to determine the size of an energy well because force can be averaged independently of which energy well the system is in, owing to the reference force always being zero. Such a feature is not naturally present in quasiharmonic analysis because a separate reference structure would be needed for every energy well. Entropy is thus divided into two types: vibrational, which relates to the size of an energy well, and topographical, which relates to the number of energy wells. The term “topographical” neutrally encapsulates related but more specific ideas about how particles are arranged in space, such as conformational or rotamer entropy for single molecules, the entropy of mixing for mixtures, cratic entropy for translational positions, orientational entropy for rotational positions, binding-pose entropy for complexes, residual entropy for crystals, the

interconversion of isomers and chemical species more generally, or even the volume per molecule in a gas. The interchanging of such terms can obscure a clear understanding of entropy.⁷⁴

The system is treated as sets of particles at different length scales: atom, united atom, molecule and system. United atom here refers to a heavy atom with all its bonded hydrogens still represented distinctly, with the whole unit moving as a rigid body, as opposed to the force-field context in which hydrogens are subsumed into the heavy atom. A molecule is similarly treated as a rigid body. United atoms and molecules each have three translational and up to three rotational degrees of freedom, depending on their symmetry. The hierarchy becomes redundant for some cases. For example, noble gases are atoms, united atoms and molecules, while water and methane are united atoms and molecules. Additional levels of hierarchy could be included for larger molecules such as polymers and proteins but these will not be considered here. The centre of a molecule is taken as its centre of mass and its axes are the principal axes, which are the eigenvectors of the moment-of-inertia tensor. The centre of the united atom is either taken as the centre of mass if it is a single molecule or as the heavy atom if it is covalently bonded to another united atom. This is because united atoms primarily pivot about the heavy atom if bonded to another united atom. The united-atom translational axes are defined as the principal axes of the molecule. The united-atom rotational axes are defined locally according to the number of bonded united atoms as follows:

- 0: The axes are the principal axes of the moment-of-inertia matrix. If there are no hydrogens, the axes are chosen arbitrarily and there are no rotational degrees of freedom.

- 1: The first axis is along the covalent bond to the united atom, the second axis is orthogonal to the plane containing the covalent bond and an arbitrarily chosen bonded hydrogen or in an arbitrary direction if there are none, and the third axis is perpendicular to both of these.
- 2: The first axis is the average of the two covalent bonds to united atoms, the second axis lies perpendicular to the first axis and in the plane of all three united atoms, and the third axis is perpendicular to both of these.
- 3: The first axis is the average of the three covalent bonds to united atoms, the second axis is defined by an arbitrarily chosen united atom of these three, and the third axis is perpendicular to both of these.
- 4: The united atom has no hydrogens, is a point and so has no rotational degrees of freedom.

For cases 1–3, if there is one hydrogen in the united atom, then there are only two rotational degrees of freedom.

The total entropy of a molecule comprises a vibrational term and a topographical term. As explained below and depicted in Figure 1, the vibrational term comprises two parts, one at the united-atom $S_{\text{vib,UA}}$, representing internal vibration, and one at the molecule level $S_{\text{vib,M}}$, representing whole-molecule motion. The topographical term also comprises two parts, one at the united-atom $S_{\text{topo,UA}}$ and one at the molecule level $S_{\text{topo,M}}$. The former is equivalent to what is commonly termed the conformational entropy; the latter comprises only an orientational term^{52–54} because the translational term, sometimes called the cratic entropy, is

zero for a pure liquid.⁵² Thus the total entropy is given by

$$S_{\text{total}} = S_{\text{vib,UA}} + S_{\text{vib,M}} + S_{\text{topo,UA}} + S_{\text{topo,M}} \quad (1)$$

and the total vibrational and topographical entropy are $S_{\text{vib}} = S_{\text{vib,UA}} + S_{\text{vib,M}}$ and

$$S_{\text{topo}} = S_{\text{topo,UA}} + S_{\text{topo,M}}.$$

Vibrational Entropy The size of each energy well is determined from the forces and torques on the particle and thermal de Broglie wavelengths and waveangles from the integrals over translational and rotational momenta.^{53,71,75} Correlations between the forces and torques of different particles are accounted for in a covariance matrix. One immediate assumption made is that the vibrational entropy of the atoms within each united atom is small and can be ignored. This is because the high-vibrational frequencies of hydrogen atoms make their contribution to entropy negligible at room temperature. A more pragmatic but not unrelated reason is that forces involving covalent bonds to hydrogens are not evaluated in the simulation because such bonds are restrained in the SHAKE protocol to allow for a larger time-step. While restraints could be dropped to allow the proper vibrational motion of hydrogens, the longer timestep necessary would not be popular with most simulation practitioners.

In greater detail, for every simulation frame, forces on each atom are first halved to account for double-counting²⁶ and rotated into the reference frame of the molecule defined by the three principal axes. Forces on each united atom i F_{xi} , F_{yi} and F_{zi} are evaluated as the vector sum of the forces of the constituent atoms. These forces are weighted by the inverse square-root of the united atom's mass m_i to give mass-weighted forces $F'_{xi} = F_{xi}/\sqrt{m_i}$, $F'_{yi} = F_{yi}/\sqrt{m_i}$ and $F'_{zi} = F_{zi}/\sqrt{m_i}$. Similarly, for united atoms, torques about each axis are

calculated from the forces and geometry, and weighted by the respective moment of inertia of the united atom to give moment-of-inertia-weighted torques $\tau'_{xi} = \tau_{xi}/\sqrt{I_{xi}}$, $\tau'_{yi} = \tau_{yi}/\sqrt{I_{yi}}$ and $\tau'_{zi} = \tau_{zi}/\sqrt{I_{zi}}$. A $6N \times 6N$ force-torque covariance matrix $\sigma_{F\tau}$ is constructed. Each element is the product of one component with another of the mass-weighted forces or moment-of-inertia-weighted torques of all united atoms in the molecule, averaged over all molecules and all simulation frames:

$$\sigma_{F\tau} = \begin{pmatrix} \langle F'_{x1} F'_{x1} \rangle & \langle F'_{x1} F'_{y1} \rangle & \langle F'_{x1} F'_{z1} \rangle & \dots & \langle F'_{x1} \tau'_{x1} \rangle & \langle F'_{x1} \tau'_{y1} \rangle & \langle F'_{x1} \tau'_{z1} \rangle & \dots \\ \langle F'_{y1} F'_{x1} \rangle & \langle F'_{y1} F'_{y1} \rangle & \langle F'_{y1} F'_{z1} \rangle & \dots & \langle F'_{y1} \tau'_{x1} \rangle & \langle F'_{y1} \tau'_{y1} \rangle & \langle F'_{y1} \tau'_{z1} \rangle & \dots \\ \langle F'_{z1} F'_{x1} \rangle & \langle F'_{z1} F'_{y1} \rangle & \langle F'_{z1} F'_{z1} \rangle & \dots & \langle F'_{z1} \tau'_{x1} \rangle & \langle F'_{z1} \tau'_{y1} \rangle & \langle F'_{z1} \tau'_{z1} \rangle & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \langle \tau'_{x1} F'_{x1} \rangle & \langle \tau'_{x1} F'_{y1} \rangle & \langle \tau'_{x1} F'_{z1} \rangle & \dots & \langle \tau'_{x1} \tau'_{x1} \rangle & \langle \tau'_{x1} \tau'_{y1} \rangle & \langle \tau'_{x1} \tau'_{z1} \rangle & \dots \\ \langle \tau'_{y1} F'_{x1} \rangle & \langle \tau'_{y1} F'_{y1} \rangle & \langle \tau'_{y1} F'_{z1} \rangle & \dots & \langle \tau'_{y1} \tau'_{x1} \rangle & \langle \tau'_{y1} \tau'_{y1} \rangle & \langle \tau'_{y1} \tau'_{z1} \rangle & \dots \\ \langle \tau'_{z1} F'_{x1} \rangle & \langle \tau'_{z1} F'_{y1} \rangle & \langle \tau'_{z1} F'_{z1} \rangle & \dots & \langle \tau'_{z1} \tau'_{x1} \rangle & \langle \tau'_{z1} \tau'_{y1} \rangle & \langle \tau'_{z1} \tau'_{z1} \rangle & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (2)$$

The mass and moment-of-inertia weighting ensures that every element in the matrix has units of J s^{-2} . Diagonalizing the matrix $\sigma_{F\tau}$ produces $6N$ eigenvalues λ_j and $6N$ eigenvectors \mathbf{v}_j .

As before,²⁹ each eigenvalue in the harmonic approximation is converted to a vibrational frequency ν_j by the equation²⁹

$$\nu_j = \frac{1}{2\pi} \sqrt{\frac{\lambda_j}{k_B T}} \quad (3)$$

where k_B is Boltzmann's constant and T is temperature. The molecule's internal vibrational entropy at the united-atom level, $S_{\text{vib,UA}}$, is calculated from the $6N - 6$ largest frequencies from the united-atom matrix, representing internal vibrations, using the equation for a quantum harmonic oscillator

$$S_{\text{vib,UA}} = R \sum_{j=1}^{6N-6} \left(\frac{h\nu_j/k_B T}{e^{h\nu_j/k_B T} - 1} - \ln(1 - e^{-h\nu_j/k_B T}) \right) \quad (4)$$

where R is the gas constant, h is Planck's constant and e is the natural logarithm base.

For whole-molecule vibrational motion, forces are evaluated from the vector sum of the united-atom forces along the three molecular axes and mass-weighted by the square root of the molecule's mass. Torques are evaluated from the united-atom forces about the three principal axes and weighted by the square root of the molecule's moment of inertia. A 6×6 force-torque covariance matrix is constructed similar to that in Eq. 2. The six eigenvalues of this matrix, corresponding to translation and rotation, are converted into vibrational frequencies using Eq. 3. The vibrational entropy at the molecular level, $S_{\text{vib},M}$, is calculated using Eq. 4 but over six degrees of freedom instead of $6N - 6$. Note that argon has no rotational degrees of freedom and so the 6 molecular degrees of freedom are replaced by 3. An inspection of the resulting eigenvectors shows that $S_{\text{vib},M}$ and $S_{\text{vib},\text{UA}}$ can be cleanly decomposed into translational and rotational entropy of the whole molecule

$$S_{\text{vib},M} = S_{\text{vib-trans},M} + S_{\text{vib-rot},M} \quad (5)$$

because three eigenvectors are found to be almost entirely force-dependent and the other three torque-dependent. While the translational frequencies were found to be very similar at both levels of hierarchy, the rotational frequencies at the united-atom level were at much higher frequency than those at the molecular level, likely because the rotational modes, being non-linear, are not cleanly extracted from linear displacements at the united-atom level.

Topographical Entropy. In a pure liquid of flexible molecules, the topographical entropy

S_{topo} has two components:

$$S_{\text{topo}} = S_{\text{topo,UA}} + S_{\text{topo,M}} \quad (6)$$

where $S_{\text{topo,UA}}$ arises from the different conformations of each molecule and $S_{\text{topo,M}}$ from the different orientations of each molecule with respect to its neighbours. $S_{\text{topo,UA}}$, reflecting the internal discretization of dihedral motion into energy wells, is evaluated from the rotamer probability distribution for all conformations at the united-atom level. The molecules considered here only have trans (t), gauche- ($g-$) and gauche+ ($g+$) conformations, defined in the usual way with boundaries at 120° , 180° and -120° . Probabilities p_i of each conformer i are evaluated from the simulation and the entropy is evaluated using

$$S_{\text{topo,UA}} = R \sum_{i=t,g-,g+} p_i \ln(1/p_i) \quad (7)$$

The number of orientations Ω_{orient} is the equivalent of the conformational term but is at the whole-molecule level relative to its external environment rather than being internal to a single molecule. It is closely equivalent to the possible multiple number of binding poses that two molecules might have when associating. To evaluate Ω_{orient} , we employ what we term the ‘‘cog-wheel’’ model^{53,75} whereby the solid rotational volume $8\pi^2$ is discretized by the volume of the neighbouring molecules. When a molecule rotates by an arc spanning the length of a neighbouring molecule, it is assumed to move into a new, equivalent energy well. Rather than explicitly evaluating volumes as in earlier work,^{53,75} which had required the awkward definition of molecular radius, we instead use the molecule’s coordination shell comprising N_c neighbouring molecules. The number of orientations Ω_{orient} depends on both N_c and the

symmetry of the molecule (see Table 1). For most molecules Ω_{orient} is given by

$$\Omega_{\text{orient}}(N_c) = \max \{1, (N_c^3 \pi)^{1/2} / \sigma\} \quad (8)$$

where σ is the molecule's symmetry number determined according to the united atoms. For every value of N_c , a molecule must have at least one orientation. The remainder of the equation is derived by assuming the coordination number is proportional to the area of a sphere by $4\pi r^2 = kN_c$, where k is a constant, and that Ω_{orient} is proportional to the area of a 3-sphere $8\pi^2 r^3 = k^{3/2} \Omega$. Solving for Ω_{orient} gives Eq. 8.

This model assumes that orientations are ideal, that is, a molecule can adopt any orientation independent of the orientation of neighbouring molecules and that each orientation has equal probability. These would appear reasonable approximations for non-polar molecules. For water, ideality is a poor approximation owing to the high cooperativity of donors and acceptors forming hydrogen bonds. Consistent with the Pauling model for ice and its generalised form for water,⁵⁴ we divide $\Omega_{\text{orient}}(N_c)$ by 4 based on the assumption that each of the two hydrogens has a 0.5 chance to donate to a neighbour, and use a symmetry number of 2. Another special case is molecules with high symmetry about one axis such as ethane and methanol. For them the number of orientations is taken as the number of surrounding molecules divided by the symmetry number:

$$\Omega_{\text{orient}}(N_c) = \max \{1, N_c / \sigma\} \quad (9)$$

For united-atom methane due to its spherical symmetry, we take $\Omega_{\text{orient}}(N_c) = 1$ for all N_c .

For all molecules, the orientational entropy is calculated for each $\Omega_{\text{orient}}(N_c)$ value, weighted

by $p(N_c)$, the probability of N_c , as

$$S_{\text{topo,M}} = R \sum_{N_c} p(N_c) \ln \Omega_{\text{orient}}(N_c) \quad (10)$$

Molecular coordination shells are defined by treating molecules as point particles and using the RAD algorithm^{76,77} which is a general way to determine molecular contacts without using fixed cutoffs. With respect to particle i , particle j is unblocked if

$$\frac{1}{r_{ij}^2} > \frac{1}{r_{ik}^2} \cos \theta_{jik} \quad (11)$$

for all particles k , where r_{ij} and r_{ik} are the distances from i to j and i to k and θ_{jik} is the angle subtended at i by j and k . Particle j contacts particle i if j and every particle closer to i than j are unblocked. Applying RAD to each molecule in every simulation frame leads to a probability distribution of molecular coordination shells $p(N_c)$ containing N_c neighbours.

Molecular Dynamics Simulations

The entropy was calculated for the liquids listed in Table 1. These are argon, water, five hydrocarbons, two arenes, three alcohols, acetone and acetic acid. For each system, molecular dynamics simulations were run using the AMBER 14 software package⁷⁸ on a cubic box of 375 identical molecules in the liquid phase. The OPLS⁷⁹ all-atom force field was used for the organic molecules, TIP4P-2005 for water,⁸⁰ and the parameters of Michels et al. for argon.⁸¹ To equilibrate each system, 500 steps of steepest-descent minimisation were carried out

followed by a 200 ps molecular dynamics simulation in the NVT ensemble using a Langevin thermostat with a time constant of 1 ps. The temperature was 298 K for all liquids except for the more volatile liquids to ensure they remained in liquid form. The temperatures used were their boiling points at ambient pressure: 87 K for argon, 112 K for methane, 184 K for ethane, 231 K for propane and 272 K for butane. Further molecular dynamics was run for 200 ps in the NPT ensemble at a pressure of 1 bar using a Berendsen barostat with a time constant of 2 ps. Coordinates and forces were saved every 10 ps in a 10 ns NPT production simulation, giving 1000 snapshots to analyse. Simulations had SHAKE on all bonds involving hydrogen atoms, a 2 fs time step, periodic boundary conditions, 8 Å cutoff and particle mesh Ewald with default AMBER parameters. Calculation of the entropy from the simulation trajectory and topology files was done using in-house C++ code.

Results

Table 2 summarises the entropies of the liquids using the method proposed here together with values from experiment⁸²⁻⁸⁴ and by the 2PT method.⁵⁸ The experimentally measured entropies of most liquids were listed on the NIST Chemistry Webbook⁸² for which a number of values have been reported by various sources. Values were obtained elsewhere for methane⁸³ and argon.⁸⁴ Errors due to accumulation of forces, torques, contacts and conformers are small and below the level of precision given. Agreement with experiment is reasonably good for most molecules, lying within the range of experiment or only a few $\text{J K}^{-1} \text{mol}^{-1}$ away. However, the entropies are worse for the three larger molecules with rings, being too large by

16 J K⁻¹ mol⁻¹ for benzene up to 44 J K⁻¹ mol⁻¹ too large for cyclohexane. Comparing to 2PT values,⁵⁸ entropies are somewhat larger for all molecules. Values here are mostly within the top of the experimental range whereas the 2PT method values fall slightly short; it should be noted that the 2PT results for water use the slightly different TIP4P-Ew force field, and other force-fields give larger values.⁵⁷ The convergence of S_{total} as a function of simulation length, from 0.01 to 10 ns, and sampling frequency, from 0.1 to 100 frames ns⁻¹ are shown in Figures 2 and 3, respectively. It can be seen that S_{total} converges rapidly. Either 0.1 ns simulation length, which is 10 frames with a sampling rate of every 10 ps, or a sampling frequency of 1 ns⁻¹, which is also 10 frames over 10 ns, are sufficient for almost all systems, except for the larger molecule toluene which requires slightly longer.

A more detailed analysis of the entropy components is presented in Table 3. $S_{\text{vib-trans,M}}$ is the largest term. The difference between $S_{\text{vib-trans,M}}$ and $S_{\text{vib-rot,M}}$ narrows for larger molecules. $S_{\text{vib,UA}}$ grows with molecule size, becoming comparable to the size of $S_{\text{vib-trans,M}}$ and $S_{\text{vib-rot,M}}$ for butane. It becomes dominant for cyclohexane and toluene, but given that their total entropy exceeds experiment, this is likely an artifact of the method. The orientational term $S_{\text{topo,M}}$ varies as a proportion of the total from as low as 5% for benzene up to 15% for propane and 16% for acetic acid. Molecules with high symmetry (Table 1) typically have lower values of $S_{\text{topo,M}}$. Using the model here augmented by a hydrogen-bond penalty, $S_{\text{topo,M}}$ of water closely matches the value of 9.9 J K⁻¹ mol⁻¹ from the generalized Pauling-model and mixing term for different numbers of hydrogen bonds.⁵⁴ Given that orientations for all other molecules are assumed to be independent of their neighbours, their $S_{\text{topo,M}}$ values are

likely to be an upper bound, especially for the systems able to form hydrogen bonds. $S_{\text{topo,UA}}$ values representing conformation are the smallest entropy component and lie slightly below the ideal value of $R \ln 3 = 9.1 \text{ J K}^{-1} \text{ mol}^{-1}$. A comparison of the entropy decomposition with 2PT reveals some similarities for the translational and rotational terms but also differences, especially in the internal term. Moreover, 2PT has no direct equivalent of $S_{\text{topo,M}}$ and $S_{\text{topo,UA}}$, instead using a diffusive component, which is included in the translational and rotational but not internal components. Evidently, entropy decomposition is non-unique, especially in the liquid-phase.

Finally, the simulation metadata to calculate entropy are examined. Figure 4 displays the force-torque covariance matrices for all liquids. These range in size from 3×3 for argon up to 36×36 for cyclohexane. The striking trend is that forces and torques are weakly correlated, as seen by the large off-diagonal blocks coloured grey. This suggests that the forces and torques could be considered in separate matrices, trading off little loss of accuracy for a smaller matrix to diagonalize. It also makes clear why considering just the self-correlations on the diagonal works so well for water²⁷ and argon.²⁶ Conversely, another not unexpected trend is that the force correlations are strongest for bonded atoms, shown by the dark bands just off diagonal. The dark colour indicates these correlations are negative. This is because the force on these atoms is dominated by the force along the bond which acts in opposing directions on each atom. Torque-torque correlations between bonded atoms are also present but weaker. For most molecules, the correlations between forces of atoms separated by more than one bond become small and negligible rapidly, the exceptions being the aromatic molecules which display greater long-range correlations that alternate in sign for each atom further away. The

6×6 molecular force-torque matrices are not displayed because they are all essentially diagonal.

Figure 5 depicts histograms of the vibrational frequencies for all molecules, converted to wavenumbers. These correspond to the so-called fingerprint region of infrared (IR) spectra. They exclude any bends or stretches due to hydrogen atoms because of the united-atom treatment that omits forces on hydrogens. Force-halving in the mean-field approximation, as discussed earlier, leads to vibrational frequencies half of that seen in IR spectra.

Table 4 contains the conformer probabilities for the two molecules with distinct conformers at the united-atom level, butane and propanol. As expected, gauche conformers are equally likely, and trans is preferred overall.

The histograms of the $p(N_c)$ distributions for all liquids are shown in Figure 6. These are all roughly Gaussian and the variation in their average value of N_c is as expected, verifying the performance of the RAD algorithm. The main contributing factor to this variation is the presence of hydrogen bonds, whose directionality brings about greater order and lower coordination, as seen for water and the smaller alcohols. The other notable deviation is the larger N_c for cyclohexane, whose non-spherical shape combined with its corrugated ring-structure prevents close packing.

Discussion

Our new method to calculate the entropy of liquids from a molecular dynamic simulation largely meets the criteria we set out for it. It is simple, making use of two covariance matrices, which are a widely used mathematical construct, albeit rarely in terms of force, the matrices' eigenvalues, and probability distributions of the coordination shells and conformations. It is informative, providing a decomposition of entropy according to normal modes relating to translational, rotational and internal vibrations with particular frequencies, as well as the number of conformations and orientations. Its convergence is rapid, with ten frames appearing sufficient for the systems studied here. This is aided by the averaging across 375 identical molecules. It is scalable to larger systems, aided by its simplicity and hierarchical nature, which lets it capture the hierarchical nature of molecular flexibility and keeps tractable the number of degrees of freedom. Further work is needed in how to demarcate an appropriate hierarchy. Its computational requirements are comparable to that of quasiharmonic analysis, although the need for matrix diagonalization with its cubic scaling on system size will slow it down and increase the need for optimization. This will be helped by the matrices' sparseness, by the possible decoupling of force and torque, and by the hierarchical treatment. It is near-general, almost able to treat all molecules equivalently, but not yet fully so because water includes a correlation term in its orientation which account for the directionality of hydrogen bonds. This necessitates including terms for the non-ideal interactions and hydrogen bonds of any molecule. Automating the generation of symmetry numbers is also required, although most molecules of interest have low symmetry. It is

accurate, reproducing experimental values for most of the liquids, the exception being the ring molecules for which it is somewhat too large in the intramolecular term. One contributing factor may be incomplete removal of whole-molecule translation and rotation from the $6N \times 6N$ matrix, leading to double-counting of entropy at the two levels of hierarchy. Another possible cause could be that force-halving to avoid double-counting²⁶ is less suitable for bonded degrees of freedom because of their stronger correlation. A third cause could be poorer convergence of off-diagonal terms in the force-torque covariance matrix for larger molecules. A fourth cause could be the harmonic approximation, which while closely adhered to based on force distributions in condensed-phase systems, will not hold in general. Such off-diagonal correlations lead to entropy that is too large,²⁹ which is the opposite to quasiharmonic analysis for which noise in off-diagonal terms brings about an entropy that is too small. Investigating this will be the subject of future investigation.

The method exhibits a number of notable features. The vibrational entropy is evaluated as an average across all energy wells, and entropy is evaluated at each structural level from a single matrix, as opposed to normal mode analysis which requires optimisation for every distinct energy well. A intriguing point of connection with other work is that Eq. 3 is equivalent to the expression for the so-called “configurational temperature”.^{33,85,86} While the theory for the orientation entropy $S_{\text{topo},M}$ may seem simplistic, it is worth emphasising that there is very little theory on discretization of rotational volume of molecules in the liquid-phase, a construction that makes treatable the high-dimensionality of configuration space e.g. 24 dimensions for a full butane-butane probability distribution. The RAD method, which had evolved from our parameter-free hydrogen-bond definition,^{54,87,88} provides a simple and

general way to determine the necessary coordination for a diverse range of molecules without any need for cutoff parameters.^{76,77} Our method’s hierarchical nature makes it applicable to simulations at different levels of atomic resolution. The use of torques also provides an efficient way to capture rotational and torsional motion, which, because of their non-linear nature, are difficult to capture in the eigenvectors of covariance matrices. Omitting torsional entropy was likely a source of the slight underestimation of entropy using united-atom force covariance.²⁹ It is not necessary to capture vibrational torsional motion directly in a conventional dihedral coordinate because of the hindering action of surrounding molecules in the condensed phase. The moment-of-inertia weighting of torque provides an effective way to average curvature over multiple configurations, rather than calculating average torques and moments-of-inertia separately as had been done previously from simulations^{52,53,75} or using the moment of inertia of a fixed geometry as is commonly done in single-molecule gas-phase entropy calculations.⁸⁹⁻⁹² Given that flexible molecules in the liquid phase are prototypes for many types of more complex system, this greatly expands the scope of entropy calculations and provides a way to quantify the full ensemble of molecular configurations. However, further developments are still required for multicomponent systems and for larger molecules and assemblies.

Conclusions

We have presented a new method to calculate the entropy of liquids of flexible molecules from a molecular dynamics simulation. The vibrational entropy component is derived from the

eigenvalues of the force-torque covariance matrices at the whole-molecule and united-atom levels, where united atoms are treated as rigid bodies rather than as point particles. The topographical entropy component, quantifying the weighted number of energy wells in the system, comprises the number of conformations and orientations. The conformational probability distributions are defined in the usual way. The number of orientations, owing to discretization of rotational volume by the surrounding molecules, in turn is derived from the molecules' distribution of coordination numbers. While further developments are planned, our method is now applicable to a much broader range of molecular systems.

Acknowledgments

We acknowledge funding from BBSRC through grant BB/K001558/1 (JH, RHH), BIOMS at Heidelberg University (RHH), and the Klaus Tschira Foundation (FG), and we are grateful to Ross Walker for implementing in AMBER 14 the capability to output force trajectories.

REFERENCES

1. C. Peter, C. Oostenbrink, A. van Dorp and W.F. van Gunsteren, *J. Chem. Phys.* **120**, 2652–2661 (2004).
2. H.X. Zhou and M.K. Gilson, *Chem. Rev.* **109**, 4092–4107 (2009).
3. H. Meirovitch, S. Cheluvarama and R.P. White, *Curr. Prot. Peptide Sci.* **10**, 229–243 (2009).
4. V. van Speybroeck, R. Gani and R.J. Meier, *Chem. Soc. Rev.* **39**, 1764–1779 (2010).
5. A.A. Polyansky, R. Zubac and B. Zagrovic, Estimation of Conformational Entropy in Protein-Ligand Interactions: A Computational Perspective, in *Computational Drug Discovery and Design*, edited by R. Baron, Vol. 819 (, , 2012), pp. 327–353.
6. R. Baron and J.A. McCammon, *Ann. Rev. Phys. Chem.* **64**, 151–175 (2013).
7. D. Suarez and N. Diaz, *Rev. Comput. Sci.* **5**, 1–26 (2015).
8. S. Kassem, M. Ahmed, S. El-Sheikh and K.H. Barakat, *J. Mol. Graph. Model.* **62**, 105–117 (2015).
9. S.H. Chong, P. Chatterjee and S. Ham, *Ann. Rev. Phys. Chem.* **68**, 117–134 (2017).
10. D.C. Wallace, *J. Chem. Phys.* **87**, 2282–2284 (1987).
11. T. Lazaridis and M. Karplus, *J. Chem. Phys.* **105**, 4294–4316 (1996).

12. O. Edholm and H.J.C. Berendsen, *Mol. Phys.* **51**, 1011–1028 (1984).
13. M.S. Head, J.A. Given and M.K. Gilson, *J. Phys. Chem. A* **101**, 1609–1618 (1997).
14. C.E. Chang, M.J. Potter and M.K. Gilson, *J. Phys. Chem. B* **107**, 1048–1055 (2003).
15. P. Vansteenkiste, V. van Speybroeck, G.B. Marin and M. Waroquier, *J. Phys. Chem. A* **107**, 3139–3145 (2003).
16. M. Tafipolsky and R. Schmid, *J. Comput. Chem.* **26**, 1579–1591 (2005).
17. B.J. Killian, J.Y. Kravitz and M.K. Gilson, *J. Chem. Phys.* **127**, 024107 (2007).
18. E. Suarez, N. Diaz and D. Suarez, *J. Chem. Theory Comput.* **7**, 2638–2653 (2011).
19. B.M. King, N.W. Silver and B. Tidor, *J. Phys. Chem. B* **116**, 2891–2904 (2012).
20. E. Suárez and D. Suárez, *J. Chem. Phys.* **137**, 084115 (2012).
21. V. Hnizdo, E. Darian, A. Fedorowicz, E. Demchuk, S. Li and H. Singh, *J. Comput. Chem.* **28**, 655–668 (2007).
22. D.J. Huggins, *J. Chem. Theory Comput.* **10**, 3617–3625 (2014).
23. U. Hensen, O.F. Lange and H. Grubmüller, *PLoS One* **5**, e9179 (2010).
24. E.B. Wilson, J.C. Decius and P.C. Cross, *Molecular Vibrations* (McGraw-Hill, London, 1955).
25. M. Karplus and J.N. Kushick, *J. Am. Chem. Soc.* **14**, 325–332 (1981).

26. R.H. Henchman, *J. Chem. Phys.* **119**, 400–406 (2003).
27. R.H. Henchman, *J. Chem. Phys.* **126**, 064504 (2007).
28. P.H. Nguyen and P. Derreumaux, *Phys. Chem. Chem. Phys.* **14**, 877–886 (2012).
29. U. Hensen, F. Gräter and R.H. Henchman, *J. Chem. Theory Comput.* **10**, 4777–4781 (2014).
30. A. Pereverzev and T.D. Sewell, *J. Chem. Phys.* **142**, 134110 (2015).
31. P.H. Berens, D.H.J. MacKay, G.M. White and K.R. Wilson, *J. Chem. Phys.* **79**, 2375–2389 (1983).
32. S.T. Lin, M. Blanco and W.A. Goddard, *J. Chem. Phys.* **119**, 11792–11805 (2003).
33. M. Martinez, M.P. Gageot, D. Borgis and R. Vuilleumier, *J. Chem. Phys.* **125**, 144106 (2006).
34. K.S. Pitzer and W.D. Gwinn, *J. Chem. Phys.* **10**, 428–440 (1942).
35. D.G. Truhlar, *J. Comput. Chem.* **12**, 266–270 (1991).
36. J. Schlitter, *Chem. Phys. Lett.* **215**, 617–621 (1993).
37. M. Klefas-Stennett and R.H. Henchman, *J. Phys. Chem. B* **112**, 3769–3776 (2008).
38. K.Y. Yeh, S.N. Huang, L.J. Chen and S.T. Lin, *J. Chem. Phys.* **141**, 044502 (2014).
39. C.E. Chang, W. Chen and M.K. Gilson, *J. Chem. Theory. Comput.* **1**, 1017–1028 (2005).

40. R. Baron, A.H. de Vries, P.H. Hunenberger and W.F. van Gunsteren, *J. Phys. Chem. B* **110**, 8464–8473 (2006).
41. O.F. Lange and H. Grubmüller, *Proteins* **62**, 1053–1061 (2006).
42. H. Schäfer, A.E. Mark and W.F. van Gunsteren, *J. Chem. Phys.* **113**, 7809–7817 (2000).
43. I. Andricioaei and M. Karplus, *J. Chem. Phys.* **115**, 6289–6292 (2001).
44. F. Reinhard and H. Grubmüller, *J. Chem. Phys.* **126**, 014102 (2007).
45. R. Baron, P.H. Hunenberger and J.A. McCammon, *J. Chem. Theo. Comput.* **5**, 3150–3160 (2009).
46. J.P. Ryckaert and A. Bellemans, *Faraday Discuss.* **66**, 95–106 (1978).
47. T.A. Weber, *J. Chem. Phys.* **69**, 2347–2354 (1978).
48. F.H. Stillinger and T.A. Weber, *Phys. Rev. A* **25**, 978–989 (1982).
49. M. Karplus, T. Ichiye and B.M. Pettitt, *Biophys. J.* **52**, 1083–1085 (1987).
50. D.J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, 2003).
51. C.E. Chang and M.K. Gilson, *J. Am. Chem. Soc.* **126**, 13156–13164 (2004).
52. S.J. Irudayam and R.H. Henchman, *J. Phys.: Condens. Matter* **22**, 284108 (2010).
53. S.J. Irudayam, R.D. Plumb and R.H. Henchman, *Faraday Discuss.* **145**, 467–485 (2010).
54. R.H. Henchman and S.J. Irudayam, *J. Phys. Chem. B* **114**, 16792–16810 (2010).

55. S.J. Irudayam and R.H. Henchman, *Mol. Phys.* **109**, 37–48 (2011).
56. G. Gyimesi, P. Zavodszky and A. Szilagyi, *J. Chem. Theory Comput.* **13**, 29–41 (2017).
57. S.T. Lin, P.K. Maiti and W.A. Goddard, *J. Phys. Chem. B* **114**, 8191–8198 (2010).
58. T.A. Pascal, S.T. Lin and W.A. Goddard, *Phys. Chem. Chem. Phys.* **13**, 169–181 (2011).
59. S. Grimme, *Chem-Eur. J.* **18**, 9955–9964 (2012).
60. S.H. Chong and S. Ham, *Chem. Phys. Lett.* **504**, 225–229 (2011).
61. H. Meirovitch, *J. Mol. Recognit.* **23**, 153–172 (2010).
62. H. Gohlke, I.Y. Ben-Shalom, H. Kopitz, S. Pfeiffer-Marek and K.H. Baringhaus, *J. Chem. Theory Chem.* **13**, 1495–1502 (2017).
63. P.K. Lai and S.T. Lin, *RSC Advances* **4**, 9522–9533 (2014).
64. T.A. Pascal and W.A. Goddard, *J. Phys. Chem. B* **118**, 5943–5956 (2014).
65. V. Hnizdo and M.K. Gilson, *Entropy* **12**, 578–590 (2010).
66. P. Durand, G. Trinquier and Y.H. Sanejouand, *Biopolymers* **34**, 759–771 (1994).
67. F. Tama, F.X. Gadea, O. Marques and Y.H. Sanejouand, *Proteins* **41**, 1–7 (2000).
68. A. Ghysels, D. van Neck, V. van Speybroeck, T. Verstraelen and M. Waroquier, *J. Chem. Phys.* **126**, 224102 (2007).

69. A. Ghysels, B.T. Miller, F.C. Pickard and B.R. Brooks, *J. Comput. Chem.* **33**, 2250–2275 (2012).
70. H. Kusumaatmaja, C.S. Whittleston and D.J. Wales, *J. Chem. Theory Comput.* **8**, 5159–5165 (2012).
71. J.A. Green, S.J. Irudayam and R.H. Henchman, *J. Chem. Thermodyn.* **43**, 868–872 (2011).
72. G. Gerogiokas, G. Calabro, R.H. Henchman, M.W.Y. Southey, R.J. Law and J. Michel, *J. Chem. Theory Comput.* **10**, 35–48 (2014).
73. L. Pauling, *J. Am. Chem. Soc.* **57**, 2680–2684 (1935).
74. A. Ben Naim, *Entropy, the Truth, the Whole Truth, and Nothing but the Truth* (World Scientific, Singapore, 2017).
75. S.J. Irudayam and R.H. Henchman, *J. Phys. Chem. B* **113**, 5871–5884 (2009).
76. J. Higham and R.H. Henchman, *J. Chem. Phys.* **145**, 084108 (2016).
77. J. Higham and R.H. Henchman, *J. Comput. Chem.* (2017).
78. D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R.

- Roe, A. Roitberg, R. Salomon-Ferrer, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, D.M. Wu, X. York and P.A. Kollman, *AMBER 2015* (, , 2015).
79. W.L. Jorgensen and J. Tirado-Rives, Proc. Natl. Acad. Sci. U.S.A. **102**, 6665–6670 (2005).
80. J.L.F. Abascal and C. Vega, J. Chem. Phys. **123**, 234505 (2005).
81. A. Michels, H. Wijker and H. Wijker, Physic A **15**, 627–633 (1949).
82. NIST Chemistry Webbook, Standard Reference Database Number 69, <http://webbook.nist.gov/chemistry/>.
83. B. Younglove and J. Ely, J. Phys. Chem. Ref. Data **16**, 577–798 (1987).
84. R.T. Jacobsen, S.G. Penoncello and E.W. Lemmon, *Thermodynamic Properties of Cryogenic Fluids* (Plenum, New York, 1997).
85. H. Rugh, Phys. Rev. Lett. **78**, 772–774 (1997).
86. B. Butler, O. Ayton, O. Jepps and D. Evans, J. Chem. Phys. **109**, 6519–6522 (1998).
87. A. Chatterjee, J. Higham and R.H. Henchman, J. Chem. Phys. **143**, 234501 (2015).
88. H. Haghighi, J. Higham and R.H. Henchman, J. Phys. Chem. B **120**, 8566–8570 (2016).
89. A.L.L. East and L. Radom, J. Chem. Phys. **106**, 6655–6674 (1997).
90. J. Gang, M.J. Pilling and S.H. Robertson, Chem. Phys. **231**, 183–192 (1998).

91. C.Y. Lin, E.I. Izgorodina and M.L. Coote, *J. Phys. Chem. A* **112**, 1956–1964 (2008).
92. J.J. Zheng and D.G. Truhlar, *J. Chem. Theory Comput.* **9**, 1356–1367 (2013).

Table 1. Molecular Symmetry Numbers

molecule	σ	number of axes
argon	-	0
water	2	3
methanol	1	2
ethanol	1	3
propanol	1	3
methane	-	0
ethane	2	2
propane	2	3
butane	2	3
cyclohexane	6	3
benzene	12	3
toluene	2	3
acetone	2	3
acetic acid	1	3

Table 2. Total Entropy Versus Experiment⁸²⁻⁸⁴ and 2PT⁵⁸ (J K⁻¹ mol⁻¹)

liquid	S_{total}	$S_{\text{experiment}}$	$S_{2\text{PT}}$
argon	53	55	
water	70	70	59
methanol	130	127-136	122
ethanol	174	160-177	141
propanol	213	192-214	
methane	78	79	
ethane	128	127	
propane	176	171	
butane	224	227-231	
cyclohexane	250	204-206	
benzene	213	197	172
toluene	259	219-221	204
acetone	211	200-221	198
acetic acid	190	158-194	147

^a Values for TIP4P-Ew model.⁵⁷

Table 3. Entropy Components versus 2PT⁵⁸ (J K⁻¹ mol⁻¹)

liquid	$S_{\text{vib-trans,M}}$		$S_{\text{vib-rot,M}}$		$S_{\text{vib,UA}}$		$S_{\text{topo,M}}$	$S_{\text{topo,UA}}$	S_{vib}	S_{topo}
	FTC	2PT	FTC	2PT	FTC	2PT				
argon	52.6		-		-		-	-	52.6	-
water	43.4	49.8	17.3	9.5	-	-	9.2	-	60.7	9.2
methanol	56.5	67.1	42.6	47.6	16.4	7.2	14.9	-	115.5	14.9
ethanol	57.1	66.7	49.6	55.1	39.8	18.9	27.7	-	146.5	27.7
propanol	59.4		52.8		62.7		29.4	8.9	174.9	38.2
methane	42.7		35.3		-		-	-	78.0	-
ethane	56.4		49.9		9.5		12.1	-	115.8	12.1
propane	63.2		58.1		29.6		25.6	-	150.9	25.6
butane	68.5		63.1		61.6		24.1	6.6	193.2	30.7
cyclohexane	69.3		67.1		94.4		19.2	-	231.0	19.2
benzene	72.5	83.0	66.2	69.6	64.1	19.8	10.4	-	202.8	10.4
toluene	72.2	82.7	65.4	73.0	96.7	47.9	24.1	-	234.4	24.1
acetone	68.2	81.0	62.4	70.5	56.6	46.4	23.7	-	187.2	23.7
acetic acid	60.8	64.8	55.9	56.0	43.5	26.3	29.7	-	160.1	29.7

Table 4. Conformer Probabilities

liquid	$p(t)$	$p(g-)$	$p(g+)$
propanol	0.712	0.142	0.146
butane	0.459	0.271	0.270

Figure Captions

Figure 1. Schematic of the four main entropy terms. The united-atom entropy includes internal molecular flexibility and the molecule entropy is whole-molecule translation and rotation. The conformational entropy is the conventional discretization into energy wells internally and the orientational entropy is the discretization of rotational volume into energy wells by the neighbouring molecules.

Figure 2. Convergence of the total entropy S_{total} as a function of time, starting from the same time origin at the end of equilibration.

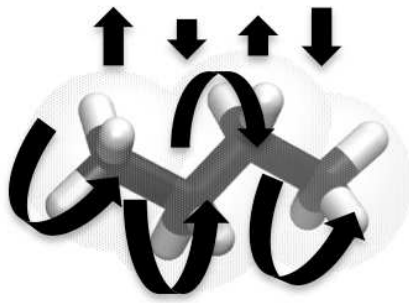
Figure 3. Convergence of the total entropy S_{total} as a function of sampling frequency over the 10 ns trajectory.

Figure 4. $6N \times 6N$ force-torque covariance matrices for all liquids, where N is the number of united atoms in the molecule. White indicates maximum positive covariance, black maximum negative covariance, and grey zero covariance. United atoms are listed for consecutively bonded atoms, starting for the alcohols from the OH end and for toluene from CH_3 , while the order for acetone is $\text{CH}_3, \text{C}, \text{CH}_3, \text{O}$ and for acetic acid is $\text{CH}_3, \text{C}, \text{OH}, \text{O}$.

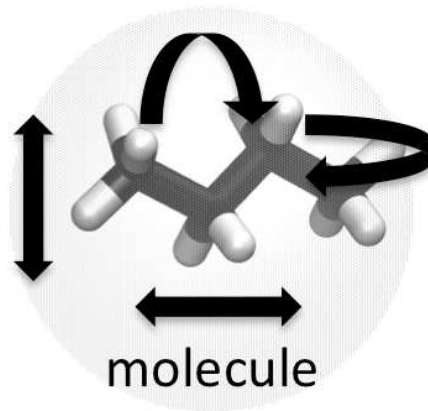
Figure 5. Wave-number ν histogram derived from the eigenvalues for each liquid. The bin-width is 50 cm^{-1} .

Figure 6. Probability distributions of coordination number N_c for all liquids.

Vibrational

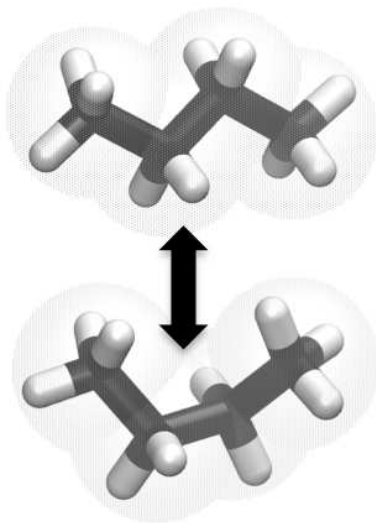


united atom

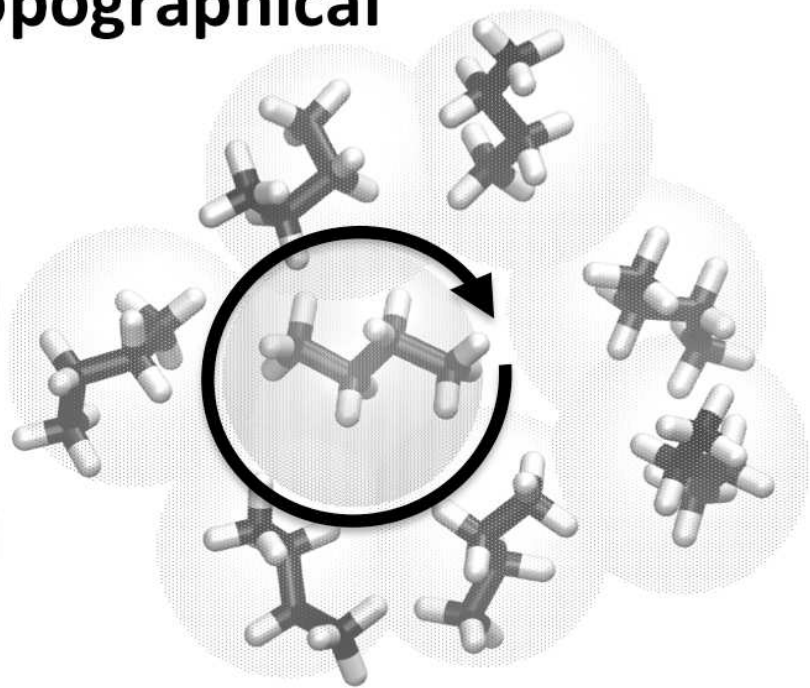


molecule

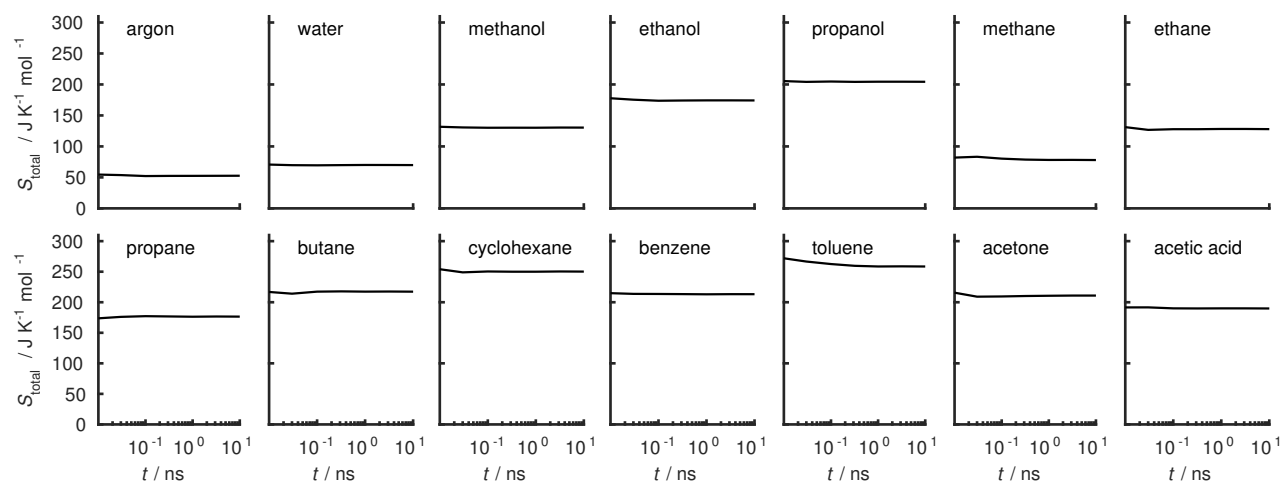
Topographical

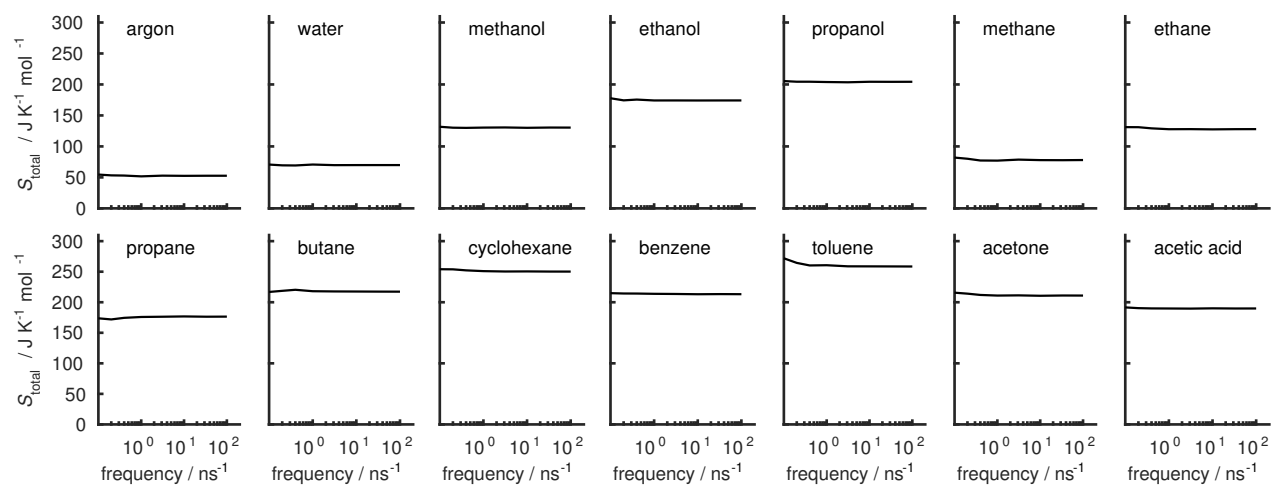


conformational



orientational





argon



water



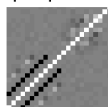
methanol



ethanol



propanol



methane



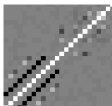
ethane



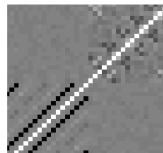
propane



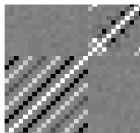
n-butane



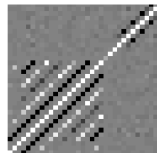
cyclohexane



benzene



toluene



acetone



acetic acid



