IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Challenges in molecular simulation of homogeneous ice nucleation

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2008 J. Phys.: Condens. Matter 20 494243

(http://iopscience.iop.org/0953-8984/20/49/494243)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 129.252.86.83
The article was downloaded on 29/05/2010 at 16:48

Please note that terms and conditions apply.

IOP Publishing Journal of Physics: Condensed Matter

J. Phys.: Condens. Matter **20** (2008) 494243 (17pp) doi:10.1088/0953-8984/20/49/494243

Challenges in molecular simulation of homogeneous ice nucleation

Andrey V Brukhno[1], Jamshed Anwar[1,3], Ruslan Davidchack[2] and Richard Handel[2]

[1] Institute of Pharmaceutical Innovation, University of Bradford, Richmond Road, Bradford BD7 1DP, UK
[2] Department of Mathematics, University of Leicester, University Road, Leicester LE1 7RH, UK

E-mail: abrukhno@gmail.com and j.anwar@bradford.ac.uk

Received 31 July 2008, in final form 25 September 2008
Published 12 November 2008
Online at stacks.iop.org/JPhysCM/20/494243

Abstract

We address the problem of recognition and growth of ice nuclei in simulation of supercooled bulk water. Bond orientation order parameters based on the spherical harmonics analysis are shown to be ineffective when applied to ice nucleation. Here we present an alternative method which robustly differentiates between hexagonal and cubic ice forms. The method is based on accumulation of the maximum projection of bond orientations onto a set of predetermined vectors, where different terms can contribute with opposite signs with the result that the irrelevant or incompatible molecular arrangements are damped out. We also introduce an *effective* cluster size by assigning a *quality weight* to each molecule in an ice-like cluster. We employ our cluster analysis in Monte Carlo simulation of homogeneous ice formation. Replica-exchange umbrella sampling is used for biasing the growth of the largest cluster and calculating the associated free energy barrier. Our results suggest that the ice formation can be seen as a two-stage process. Initially, short tetrahedrally arranged threads and rings are present; these become correlated and form a diffuse ice-genic network. Later, hydrogen bond arrangements within the amorphous ice-like structure gradually settle down and simultaneously 'tune-up' nearby water molecules. As a result, a well-shaped ice core emerges and spreads throughout the system. The process is very slow and diverse owing to the rough energetic landscape and sluggish molecular motion in supercooled water, while large configurational fluctuations are needed for crystallization to occur. In the small systems studied so far the highly cooperative molecular rearrangements eventually lead to a relatively fast percolation of the forming ice structure through the periodic boundaries, which inevitably affects the simulation results.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Water, while being one of the simplest and ubiquitous compounds on Earth and providing the environment for a vast diversity of phenomena including life itself, has long presented challenges in understanding its peculiar properties and complex actions. Indeed, it is remarkable how such a simple molecule, when present in large numbers, can give rise to so many different emergent manifestations. The chameleonic behaviour of water is also well reflected by the fact that under different external conditions it can solidify in over a dozen distinct ice polymorphs [1, 2], two of which—hexagonal (I_h) and cubic (I_c)—can occur at the same time individually and in mixed structures under ambient pressure. The spread and importance of the two ice forms I_h and I_c cannot be overestimated, considering that the polar ice caps, which mostly consist of these two ices, contain about 80% of fresh water on Earth,—a huge and invaluable reserve on which the planet's ecology and our survival depends. In fact,

[3] Author to whom any correspondence should be addressed.

 © 2008 IOP Publishing Ltd Printed in the UK

crystal growth in polar ice is indicative of long-term climate changes [3, 4], whereas understanding of ice formation in clouds may help in analysing rain- and snow-fall patterns as well as in estimating the intensity of solar radiation passing through the atmosphere to Earth's surface [5, 6]. In everyday life we mostly encounter hexagonal ice I_h (possibly mixed with I_c), known to us in the shape of frost, snow, icicles and solid ice surfaces. Also, it is the ambient ice that embraces tissues and other ingredients of frozen food and, therefore, bears some responsibility for the robustness of food preservation, which is a mild example of the issues associated with biological tissue preservation [7].

While water is well known for its exclusive vital role in evolution and life support, ice, in its turn, is believed to be the primary candidate for the role of a molecular container suitable for conservation and transmission of life forms through outer space,—sort of a natural, molecular level counterpart of Noah's ark. This is one of the reasons for the international effort in finding water/ice patches and traces on planets and in other space objects such as comets, meteorites and asteroids [8, 9]. Finally, there is a recent hypothesis that initial steps of RNA evolution may take place very slowly within the permafrost on the timescale of longlasting ice ages [10]. In this long-term process local stresses and molecular rearrangements within the subsoil ice, due to weather fluctuations and palaeoplatform drifts, should be a major factor leading to a retarded diffusion of icebound small organic compounds (methane and nitrates) which are supposed to eventually meet and react.

That said, understanding and quantifying crystallization of water into different ice phases under various external, as well as unaided, conditions is, without doubt, of significant scientific and technological importance. However, it appears that this understanding, particularly the early stages of nucleation, is rather difficult to investigate experimentally [11, 12], as the required spatial and temporal resolution is not quite accessible by experiment. In such circumstances molecular simulation can serve as a powerful alternative means for studying the evolution of local molecular structure and for estimating thermodynamic quantities such as free energy barriers. Indeed, during the last decade, ice nucleation and homogeneous crystallization in bulk water has become accessible to simulation [13–18]. Several studies of spontaneous ice formation in water have been reported, where either small ice nuclei were directly observed in simulation [14] or rare molecular dynamics (MD) trajectories connecting liquid water and crystal ice phases were generated [15, 16]. Also, subsurface freezing of neat and salty water has been extensively studied in MD simulations, being supported by high-speed visual (VIS) and infra-red (IR) imaging [17]. However, as has been stressed already by Matsumoto *et al* [15] and also by Radhakrishnan and Trout [18], a satisfactory description of the nucleation process can only come from observations of a sufficiently large ensemble of trajectories, or paths, leading from the metastable liquid to stable crystal phase and vice versa. Of course, determining the corresponding free energy barrier(s) associated with the formation of the critical nuclei and possible intermediate cluster structures would be invaluable in identifying the relative importance of various pathways of crystallization [18, 19].

1.1. Issues of ergodicity

A rigorous study of the homogeneous liquid–solid transition by computer simulation is particularly difficult for two main reasons: (1) the nucleation process is relatively sluggish with a very small window in terms of optimum supercooling, and (2) numerous diverse pathways can characterize the nucleation step [14, 15, 20–24]. Homogeneous ice nucleation is known to be extremely hindered in comparison to its heterogeneous analogue (ice formation on a substrate, or at an interface), the rate of the former being estimated as about 10^{10} times slower than that of the latter [5]. The hindrance can be partly ascribed to the enormous entropy loss by water molecules when the ice structure is formed, so that tiny entropic bottlenecks (better to say 'needle-eyes') must be passed prior to the formation of a noteworthy nucleus. Moreover, different nucleation paths are likely to be distant and well separated from each other in the configuration space owing to the diversity of possible water structure patterns and the associated *severe roughness of the energetic landscape*. This results in intrinsic *intermittent dynamics*,—typical of so-called frustrated (glass-like) systems [15]. Some of the paths are, of course, expected to be easier to follow than others. Yet, energetically the most preferable (the easiest) trajectories are considered to constitute only a small subset of the entire path ensemble available–the situation common with the protein folding phenomenon. That is, in the absence of interfaces or nucleating agents, a simulated *finite* system may struggle to find favourable pathways to nucleation, so that the contribution of these paths to the actual crystallization process may, in fact, appear faded. This is how the *path entropy* and the related issues of ergodicity due to both *entropic and energetic bottlenecks* enter the simulation problem. Evidently, linked to these is a general effect of progressively damped kinetics as the rate of supercooling is increased, effectively leaving only a small window for optimum supercooling. Altogether this implies that (i) extremely long simulations are necessary for bulk water–ice systems and (ii) extraordinary care is to be taken to ensure that the longest relaxation times are overcome, so as to allow for the most important regions in the configuration space to be sampled. These two propositions are well supported by both the earlier simulation studies [14–17] and our own experience in simulation of the water-to-ice transition.

In the above discussion we were mostly referring to brute-force molecular dynamics simulations of spontaneous ice formation in water, for which the simulation times need to be comparable with the real world nucleation times (microseconds). There are, of course, smarter ways of tackling the problem, for example by effectively directing the simulation towards and over the barriers that otherwise represent great obstacles. One of the most efficient means of doing so is the umbrella sampling (UMS) method [26, 27]. Although there have been reported seemingly faster and more robust approaches which make use of self-consistently converging penalty functions [28–32], from our experience [33] we have learnt that these become progressively error-prone for systems where the dynamics is considerably damped by strong short-range correlations. Supercooled fluids and, in particular, water at low temperatures are typical

examples of such systems. In fact, umbrella sampling shows its full strength only when it is carried out in sufficiently small so-called windows and especially when it is combined with a replica-exchange scheme. This way several studies of nucleation and crystallization of Lennard-Jones type systems have been performed [20–23], as well as crystallization of distinct ice phases from supercooled water [18].

1.2. Issues of order parameters

In order to direct a water–ice phase transition in a simulation one needs to develop an appropriate order parameter that quantifies the degree of crystallization by mathematically discriminating between liquid- and solid-like local arrangements around a particle/molecule. In studies where the simulated substance has been modelled by spherical particles [20–23], orientation order parameters were based on spherical harmonics analysis [34]. However, when applied to the water–ice phase transition, this approach appears to be inconclusive and unusable unless additional orientational order measures, such as tetrahedral angular correlation [35], are included [18, 19]. Indeed, recognition of crystallogenic structures in water is more involved because, unlike with spherical particles, the local order in ice structures cannot be characterized in terms of plain site-to-site similarities in the orientational patterns of neighbouring water molecules, at least not in the way spherical harmonics are commonly used. The tetrahedral symmetry in the positions of oxygen (O-) atoms within the first coordination shell (FCS) of each molecule, while being typical of ice crystals, serves as an additional hint only, because the local H-bonding in water is also largely close to being tetrahedral. Furthermore, for the seemingly same tetrahedral order to propagate beyond the FCS, the tetrahedra in each next layer should be arranged so as to share their links with both the inner and outer neighbouring tetrahedra. Apparently, several distinct tetrahedral lattices can emerge, as is revealed by the variety of the known ice polymorphs [2]. In certain cases these can be further superimposed onto each other—the vertices of one coinciding with the centres of the other. For instance, ice VII/VIII formed from liquid water at pressure above 3 GPa can be seen as two interpenetrating I_c lattices. To give another example of combined ice structures, the basal-plane sheets of normal ices, I_h and I_c, make a perfect match and can alternate in arbitrary order and number [36, 37]. Thus, it is the secondary, long-range structure (plus H-bond order) that actually determines a particular ice phase and, hence, adds to the complexity that must be dealt with.

1.3. Our aim: competitive growth of ices I_h and I_c

We noted earlier the possibility that the nucleation process may occur via a diverse series of pathways. This issue is further compounded when two or more structural phases can emerge in a sequence or otherwise, as it occurs in simple fluids where a metastable body-centred crystal (bcc) nucleus forms first and only later the stable face-centred crystal (fcc) structure develops and spreads over the volume [20–23]. A similar phenomenon of phase switching can be observed in binary mixtures of hard spheres as well [38].

Even though experimental evidence [39] and earlier simulations [18] strongly indicate that it is hexagonal ice which is stable at temperatures above 160 K with cubic ice being more stable below this threshold, very little is known about the details of crystallization in bulk water. There is also some evidence that ice I_c is favoured in the upper atmosphere [5, 6] and in water droplets smaller than 30 nm at $T = 160$–220 K [40, 41]. The latter studies as well as simulations of crystallization of pure ice phases [18, 19] suggest that cubic ice has a lower barrier to formation than hexagonal ice owing to its lower interfacial free energy. This motivates our own research aimed at modelling homogeneous ice nucleation without restrictions on the ice phase to grow.

In this report we focus on the development of appropriate order parameters for recognizing and directing the growth of *competing* ice phases, I_h and I_c, i.e. with no bias favouring any particular form, in supercooled water simulated under ambient pressure, 0.1 MPa. We first revisit the use of the spherical harmonics in orientational order characterization (section 2), and highlight the inconsistencies in the description of the two ice structures when using the harmonic Y_{32}—the only harmonic with tetrahedral geometry, and hence seemingly the most suitable for ices. We show that applicability of the harmonics is restricted to the cubic ice structure only, whereas the orientational pattern characteristic of hexagonal ice cannot be fully described. In section 3, we introduce a maximum projection method which is general and flexible enough for *simultaneous* search of arbitrary, albeit predetermined, geometries. Based on this, we define orientational order parameters that can robustly differentiate between low-structured water and two distinct molecular configurations that characterize I_h and I_c lattices. The proposed scheme allows us to grow both ice structures competitively at the same time, with the possibility of yielding mixed structures. It also enables a clear distinction to be made between the two crystal patterns. Furthermore, our study shows that biasing the nucleation process is more efficient when done with respect to an *effective* cluster size rather than an absolute cluster size comprising a discrete number of molecules identified as belonging to the same cluster. That is, we assign a *quality weight* to each molecule within apparently ice-like neighbourhood so as to damp out the contribution of molecules that are connected but aligned poorly. This results in the biasing procedure becoming notably more sensitive to gradual improvements in the sought molecular order within the nucleus. In section 4, we present preliminary results from our Monte Carlo simulations of homogeneous ice nucleation, where both ice structures, I_h and I_c, emerge in the course of nucleation and crystal growth.

2. Spherical harmonics analysis revisited: issues with ice

We begin with the terminology widely used in application of spherical harmonics to crystal structure recognition during simulations of liquid–solid transitions. The first coordination shell of a particle is defined as its vicinity within a sphere of radius $R_{\min}$, corresponding to the first minimum in the pair radial distribution function (≈3.45 Å for water/ice). All

particles (j) within the first coordination shell of a given particle (i) are called *neighbours*, $r_{ij} < R_{\min}$, whereas the corresponding radii-vectors $\mathbf{r}_{ij}$ are called *bonds*. The latter should be distinguished from (solid-like) *connections* which are subject to additional criteria imposed on the bond correlations of two neighbouring particles.

Spherical harmonics analysis is based on a angular analogue of the Fourier series used for approximating an arbitrary function of one variable. That is, in principle, any (angular) function defined on a unit sphere can be expanded in a series of orthogonal complex functions, $Y_{lm}(\theta, \phi)$, the so-called spherical harmonics, where θ and ϕ are the azimuthal and polar angles in spherical coordinates, l is a positive integer, and integer $m \in [-l, l]$. Evidently, $r = 1$ on the unit sphere, and in general the radial dependence can be introduced by multiplying the series by a radial function.

2.1. The standard approach to recognition and biasing of the crystal growth

The method of spherical harmonics for capturing and analysing cluster structures in condensed matter was first proposed by Steinhardt *et al* [34] and has been utilized effectively in a number of nucleation and crystallization studies [18–23]. The essential idea is to estimate the (probability) density of so-called bond orientations, $\rho(\theta, \phi)$, between a probe particle and its nearest neighbours by expansion in spherical harmonics:

$$\rho(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \bar{q}_{lm} Y^*_{lm}, \tag{1}$$

where $\bar{q}_{lm}$ denotes the overlap between $\rho(\theta, \phi)$ and Y_{lm}. In a fully disordered, isotropic substance, such as gas or liquid, $\rho(\theta, \phi)$ is symmetric and uniform, which implies that the only non-vanishing contribution in (1) is due to $l = 0$. However, when crystallization occurs, a well-pronounced orientational order emerges at the crystal nucleation centres. In such circumstances some of the coefficients $\bar{q}_{lm}$ with $l \neq 0$ become non-zero, and by using those it is possible to measure the *average, or global, orientational order* in the system. Note that, in practice, the expansion has to be cut at some finite value of l, and the more terms included in (1) the better the approximation of $\rho(\theta, \phi)$ obtained.

From the perspective of cluster analysis in simulations, it is actually more useful to introduce an instantaneous local counterpart of $\bar{q}_{lm}$. That is, bond orientation order around particle i can be quantified by a set of complex numbers,

$$q^{(i)}_{lm} = \frac{1}{N^{(i)}_{\rm b}} \sum_{j}^{N^{(i)}_{\rm b}} Y_{lm}(\theta(\mathbf{r}_{ij}), \phi(\mathbf{r}_{ij})), \tag{2}$$

where the azimuthal and polar angles, θ and ϕ, determine the orientation of the bond, $\mathbf{r}_{ij}$, between particles i and j, and $N^{(i)}_{\rm b}$ is the number of bonds for particle i. Clearly, $q^{(i)}_{lm}$ is not invariant with respect to rotation of the coordinate frame. For a particular value of l, though, these $(2l + 1)$ numbers can be viewed as components of a complex vector, $\mathbf{q}_l$, the norm of which being, of course, invariant:

$$q^{(i)}_{l} = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^{l} q^{(i)}_{lm} q^{*(i)}_{lm} \right)^{\frac{1}{2}}. \tag{3}$$

In a dilute phase where particle movements are correlated weakly, the bond orientations around different particles are distributed randomly, and thus, the distribution of norms, (3), is such that their maxima are close to zero. As the particle correlations become stronger, e.g. in liquids close to freezing and especially upon supercooling, certain local order gradually arises, thereby breaking the local symmetry of bond orientations and shifting the corresponding q_l distributions away from zero. Yet, the vectors $\mathbf{q}_l$ for different particles remain mostly uncorrelated owing to the overall isotropy in liquids. This is in contrast to well-ordered crystal structures, where bond orientations are strongly correlated which, in its turn, implies collinearity of $\mathbf{q}_l$ vectors with the same l-value. Also, for a perfect crystal, the magnitude of vectors *characteristic of the crystal structure* is at maximum.

In a crystal at non-zero temperature the structural order is of course to some degree diffuse, and $\mathbf{q}_l$ vectors of different particles are never completely collinear. However, the correlation between two vectors, for particles i and j, can serve as a convenient measure of the similarity between the nearby surroundings of the two particles,

$$q^{(ij)}_{l,\rm corr} \equiv \frac{\mathbf{q}^{(i)}_l \mathbf{q}^{*(j)}_l}{q^{(i)}_l q^{(j)}_l} \quad (\to 1 \text{ in perfect alignment}). \tag{4}$$

Indeed, the criterium of $q^{(ij)}_{l,\rm corr} > \tau_{\rm corr}$ is often employed to identify solid-like connections between neighbouring particles [20, 21], $\tau_{\rm corr}$ being the correlation threshold value. Note that (i) $q^{(ij)}_{l,\rm corr}$ is also invariant under coordinate rotation, and (ii) even when the norms, $q^{(i)}_l$, for a given l are small (i.e. Y_{lm} is not characteristic of the structure) the correlation criterium is still valid, although it may be insufficient and, for that reason, inefficient.

By taking the average in (2) over the entire system we arrive at the definition of the corresponding measures of global order,

$$Q_{lm} = \frac{1}{N} \sum_{i}^{N} N^{(i)}_{\rm b} q^{(i)}_{lm} \tag{5}$$

which coincide with the coefficients in (1) in the limit of $N \to \infty$. The global invariants are then determined as,

$$Q_l = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^{l} Q_{lm} Q^*_{lm} \right)^{\frac{1}{2}}. \tag{6}$$

Note that in (5) the radius-vector of each bond, $\mathbf{r}_{ij}$, is accounted for twice (once for atom i and secondly for atom j), but with opposite signs. This double-counting makes zero all Q_{lm} vectors with odd l-values, because the corresponding harmonics are antisymmetric.

The definitions (1)–(6) serve to provide the required theoretical basis and a direct prescription as to how to monitor

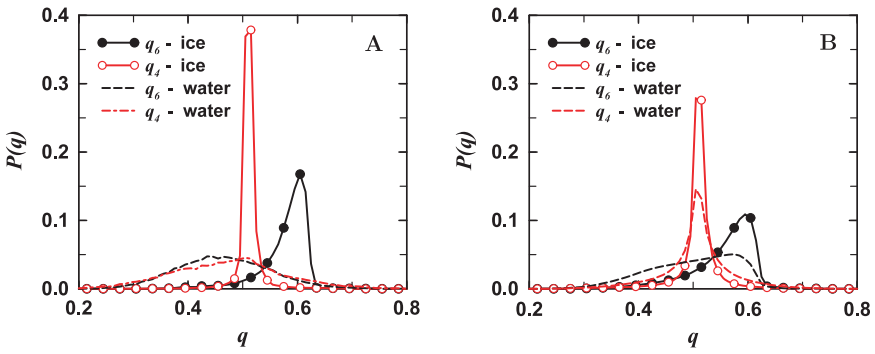

Figure 1. The probability distributions of q_4 and q_6 obtained with TIP4P water model for ice I_h and water. (A) Stable phases at $T = 150$ and 300 K, respectively; (B) supercooled water and stable ice at 200 K $< T_{melt} = 230$ K.

the orientation order, both on the local and global scales. In practical simulations, however, including harmonics of several orders at once is computationally too expensive. Thus, it is important to make a reasonable selection of l so that only the relevant harmonics are accounted for. Therefore it has become a common practice to include into the analysis only the harmonics with $l = 4$ and 6 which seem to have reasonably diverse shapes and indeed have proven to capture well simple cubic (scc), face-centred (fcc), body-centred (bcc) and hexagonal (hcc) crystal structures found in dense systems of spherical particles [20–23]. In particular, harmonics with $l = 6$ appear to be less sensitive to the differences between various phases, which is the reason for q_{6m} and Q_6 being usually used as common, correspondingly local and global, order parameters. Then q_{4m} and Q_4 are used only for a more detailed analysis, such as discrimination between crystal phases.

Clearly, the local and global parameters are not equivalent. The former, by definition, measures orientational order around a particular particle/molecule and, hence, provides a convenient means for tracking the process of nucleation locally. Equation (4) is then employed as a criterium for establishing solid-like connections between particles [20, 21]. All neighbouring particles having more than a certain number of connections, $N_{cn} > N_\tau$, are included into a solid-like cluster, i.e. a crystal nucleus. In contrast, global parameters measure only average orientational order while ignoring all the details of local clustering. For instance, in LJ-type systems nucleation proceeds stepwise—first, small bcc nuclei are formed, and then within a sufficiently large bcc nucleus an fcc core emerges; eventually the fcc structure, being the actual stable phase, spreads over the entire volume [20–23]. By using locally defined parameters it is possible to find all separate nuclei and identify all phase patterns, whereby gaining quite detailed structural information. One can, of course, use the parameters of both types in the same simulation [19, 20].

An important property of a good orientational order parameter is that it should map onto the required physical structure without being too degenerate. However, in general, we aspire to have an order parameter that yields a solid (in contrast to the liquid) state without preference for any particular crystalline structure that may emerge. In view of this conflict the specification for the order parameter is not always obvious *a priori*. The latter issue becomes especially pronounced in molecular fluids such as water. For instance, considering low symmetry and, hence, high anisotropy of molecular crystals, the rotational invariance of an order parameter may be associated with undesirable degeneracy with respect to orientation of small crystal-like molecular arrangements. In practice, applying a rotationally degenerate bias can lead to fixating of incoherently grown, albeit locally correct, patterns so that the formation of the long-range crystal lattice would be unwittingly hindered. As we discussed in section 1.2, and will see below, this appears to be part of the issue in application of the spherical harmonics to directing the growth of normal (I_h and I_c) ice structures in water.

It is important to realize, though, that rotational invariance is *not a necessary* requirement for an order parameter to be robust and applicable for *biasing* the nucleation process (as contrasted to *monitoring* the crystal growth statistics where a rotationally invariant measure is actually needed so as to detect differently oriented nuclei). Indeed, as long as the parent phase is isotropic (liquids are such), in a sufficiently large volume nucleation centres of any possible orientation can be found. Therefore, the imposed orientation of the crystal nucleus to be grown can be chosen arbitrarily, e.g. solely for the convenience of simulation, without any effect on the estimated barrier shape.

2.2. *Inconsistency between the normal ice structures and spherical harmonics*

One striking example illustrating a failure with the harmonics Y_{4m} and Y_{6m} has been reported in the case of water–ice transformation [18]. The sole use of Q_4 and/or Q_6 was insufficient to direct a simulation starting in supercooled water towards the ice I_h crystal. A clue to this failure can be found in figure 1 where the q_4 and q_6 distributions obtained with TIP4P water model are shown for equilibrated water and ice I_h phases ($T = 300$ and 150 K, respectively) as well as for metastable supercooled water and stable ice at the same temperature, $T = 200$ K $< T_{melt} = 230$ K [1, 25]. We see that, even though both q_4 and q_6 distributions exhibit marked difference in their shape and maxima locations between the stable phases, the differences virtually disappear when comparing supercooled water and ice crystal simulated below T_{melt}. Moreover, there is

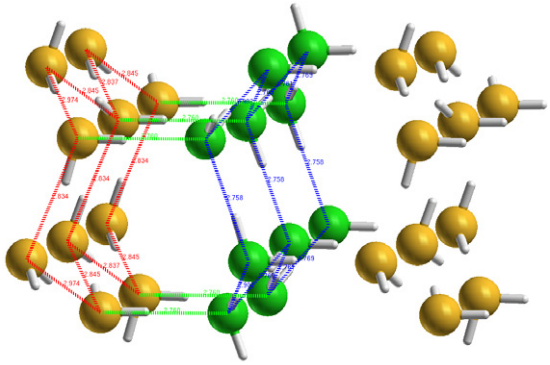

Figure 2. Illustration of the alternating mirror-image clusters distinguished within a single crystal structure of ice I_h by using $\mathbf{q}_6$; see (4) and the text for more details. H-atoms are skipped for clarity. The mis-split basal sheets are depicted in different tones. The horizontal OO-connections are missed by $\mathbf{q}_6$-products.

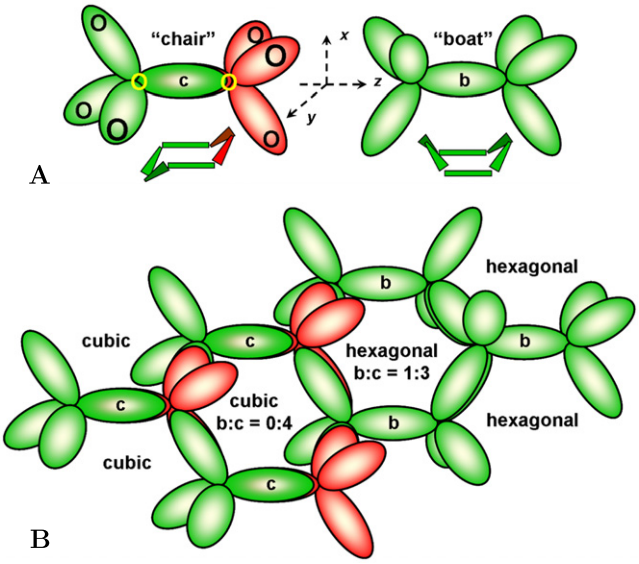

Figure 3. (A) The two configurations of eight water molecules found in normal ice structures, I_h and I_c, and mixtures thereof. The angular diffuseness of OO-connections (H-bonds) around the links in the ideal ice lattice(s) is depicted in a fashion conventional for spherical harmonics—by asymmetric dumbbells, or lobes. (B) A sketch of mixed ice structure. Cubic phase contains staggered, or 'chair' (c), configurations only, whereas eclipsed, or 'boat' (b), configurations are found in hexagonal phase, in 1:3 proportion to 'chair'-forms.

another fundamental issue. The cluster search procedure based on finding correlations between $\mathbf{q}_6$-vectors, (4), when applied to the perfect ice I_h structure, mistakes the entire single crystal for alternating (unconnected) sheet-like clusters which are, in fact, mirror images of each other, see figure 2. Apparently, this mis-splitting is provoked by the conflicting symmetries of two local molecular configurations present in ice I_h, only one of which being consistent with the symmetry properties of Y_{6m}. That is, in contrast to ice I_c where each water molecule has all its four H-bonds involved in a *staggered* structure, one of the H-bonds per molecule in ice I_h participates in an *eclipsed* arrangement, whereas the other three H-bonds are found in the same staggered configuration as in ice I_c, see figure 3(A). Interestingly, in a perfect I_h crystal all the H-bonds in the eclipsed formation are aligned perpendicular to the basal plane (along c-axis of the crystal, z-axis in figure 3), and these are just the OO-connections missing in figure 2 where only O-atoms within the same vertical sheet have been identified as interconnected (depicted in the same colour).

Let us look closely at the issue and the challenge that it creates. The H-bonding or the equivalent OO-connections give rise to two distinct supramolecular arrangements, namely the so-called 'chair' (staggered) and 'boat' (eclipsed) configuration which combine to give the hexamer rings, see figure 3(A). It is easy to see that all such OO-configurations that are not aligned with the c-axis of the crystal (z-axis in figure 3) are of the chair form, whilst those that are aligned with the c-axis can adopt either chair- or boat-like configuration, thereby defining the type of ice structure, I_c or I_h, accordingly. It is also clear that the two ice phases can have a coherent interface and, in principle, mixed phases can occur by alternating the configuration adopted in the basal-plane sheets along the z-axis in figure 3. The two supramolecular arrangements, chair and boat forms, have different symmetry properties—the former being *antisymmetric* whereas the latter having *even symmetry* with respect to the positions of the two O-atoms making up a connection (upon superposing these O-atoms). Note that in either case the H-bond tetrahedral patterns around the two neighbouring O-atoms cannot be translated so as to coincide, i.e. one of the tetrahedra has to be re-oriented to enable coincidence. Therefore, instead of looking for similarities in the neighbourhood of two molecules (which is sufficient for spherical particles), the search for ice-like orientational order in water should be based on a more specific criteria, marking as 'ice-like' only the two relevant molecular arrangements while disregarding all other possible patterns.

To exemplify, we can try to fit the two ice-genic configurations with Y_{32}—the only harmonic possessing tetrahedral symmetry and, thus, being most relevant for ices. The result of such an attempt is shown in figure 4. $\mathrm{Re}(Y_{32})$ set represents two embedded tetrahedra inverse with respect to each other and taken with opposite signs, i.e. antisymmetric. All chair-form configurations present in ice can be mapped onto these two tetrahedra, as long as one of the lobes in $\mathrm{Re}(Y_{32})$ is aligned with the crystal c-axis. Note how, due to antisymmetry, the corresponding q_{32}-product should be taken with a minus sign in order to produce a positive correlation measure between the inverse tetrahedra making up a chair-form configuration, see figure 4(A). Clearly, in order to construct (or distinguish) a boat-form configuration, *one of the off-axis triplets* in $\mathrm{Re}(Y_{32})$ has to be rotated 60° about the crystal c-axis, as illustrated in figure 4(B). One might think that collecting overlaps with these harmonics—$\mathrm{Re}(Y_{32})$ and its rotated version $\mathrm{Re}(Y_{32})^{\mathrm{rot}}$—in a single $\mathbf{q}$-vector could resolve the issue. An additional complication arises, though, because in ice for any pair of neighbouring water molecules (O-atoms), the two arrangement types are mutually exclusive, i.e. only one of them—either chair or boat form—can be assigned to a particular OO-connection. In contrast, in supercooled water H-bonding patterns often produce diffuse

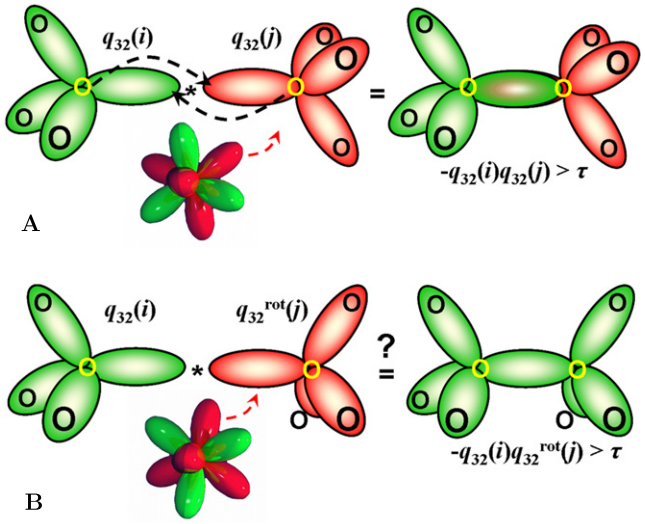

Figure 4. (A) Construction of q_{32}-product for two water molecules sharing an OO-connection (H-bond) in ice, which illustrates the *antisymmetric correlation* between the two tetrahedra forming the chair-like (staggered) configuration. The underlying spherical harmonics, $\mathrm{Re}(Y_{32})$, is shown below the sketch. (B) Illustration of 60° rotation adjustment necessary for one of the tetrahedra within $\mathrm{Re}(Y_{32})$ ((red) dumbbell triplet) in order to obtain the correct correlation product for the boat-form (eclipsed) configuration. The incompatibility of the chair and boat arrangements is highlighted by questioning the necessity of sign (shading) switch upon rotation of $\mathrm{Re}(Y_{32})$.

configurations that can partly meet both the boat and chair criteria simultaneously. For this reason the two forms (A and B in figure 4) must be counted with opposite signs so as to damp out any mixed arrangements—the condition that overrides the inherent antisymmetry of Y_{32}. The implication is that one cannot combine $\mathrm{Re}(Y_{32})$ with its rotated version and have a unified criterium for the correlation products describing mutually exclusive events. Indeed, irrespective of the harmonics used, as long as symmetric and antisymmetric molecular arrangements are allowed within the same structure, it seems impossible to unambiguously distinguish both types by imposing a single condition on the dot-product between **q**-vectors.

Our pictorial analysis shows why the property of rotational invariance is irrelevant in this case, whereby such global parameters as Q_4 and Q_6 or a simple tetrahedral correlation measure [18, 19, 35], are rotationally too degenerate for effectively biasing the growth of ice(s) I. It comes as no surprise that collinearity of $\mathbf{q}_l$-vectors, (4), cannot fully describe structural connectivity between water molecules within the ambient ice forms.

3. Maximum projection method

Given that a simple cluster analysis based on the use of spherical harmonics is not deemed adequate to direct ice formation, we now introduce an alternative procedure, which does not suffer from the aforementioned shortcomings and is sufficiently flexible to be adapted for other analogous problems.

We need a measure of orientation correlations with respect to a particular set of directions. As with spherical harmonics, certain spots on a unit sphere should be highlighted whereas the contribution of intermediate regions should vanish. Such a measure can be based on angular correlations between a vector in question, i.e. a given bond, and a set of predefined vectors of interest, or *directors*. In the most straightforward implementation one could merely collect a normalized sum of directing cosines as it has been done earlier in the case of nitromethane [42]. Below we generalize the latter approach so as to be able to include a few—in our case two—mutually exclusive director configurations.

Let us consider a set of directors, $\{\mathbf{d}_k\}$, each starting at the origin (on the O-atom of a water molecule in focus) and pointing to one of the vertices of a tetrahedron. For certainty, we assume one of the four directors being aligned with the z-axis ($x = 0$, $y = 0$) and one of the remaining three placed in the xz-plane ($y = 0$). In order to account for all distinct orientations of the main tetrahedron we extend our director set by including vectors with all possible sign variations:

$$\begin{aligned} \mathbf{d}_1 &= (0, & 0, & & \pm 1), \\ \mathbf{d}_2 &= (\pm \sin\gamma, & 0, & & \pm\cos\gamma), \\ \mathbf{d}_3 &= (\pm \sin\gamma \cos\tfrac{\pi}{3}, & \pm \sin\gamma \sin\tfrac{\pi}{3}, & & \pm\cos\gamma), \end{aligned} \qquad (7)$$

where γ is the tetrahedral angle $\cos\gamma = -\frac{1}{3}$. Thus, the total set consists of 14 directors being counterparts of the lobes of $\mathrm{Re}(Y_{32})$ and its rotated version shown in figure 4. However not all of the included tetrahedral sets can be paired in accordance with the two ice-genic configurations. Appropriate pairing of the tetrahedra for two neighbouring water molecules is introduced by using the following projection summation procedure.

For a given (arbitrary) unit vector, **u**, one can always find the director of maximum correlation, i.e. the one corresponding to $\max\{\mathbf{u}\mathbf{d}_k\} = \cos\psi_{\min}^{\{k\}}(\mathbf{u})$, where $\psi_{\min}^{\{k\}}(\mathbf{u})$ is the actual projection angle (which is minimum, of course). Then, in a fashion similar to the spherical harmonics analysis, for a number of normalized vectors, $\{\mathbf{u}_{ij}\}$, representing bonds around the molecule in focus (ith) one can *in general* define a local orientation measure specific for the given director set,

$$q_{\mathrm{dir}}^{(i)} = \frac{1}{N_{\mathrm{b}}^{(i)}} \sum_{j}^{N_{\mathrm{b}}^{(i)}} (\pm)[\max_{k}(\mathbf{u}_{ij}\mathbf{d}_k)]^n. \qquad (8)$$

Here we introduced the (yet undefined) signs within the sum in order to impose selective correlation rules for damping out, where appropriate, the contributions inconsistent with the dominant geometry of connections, figure 5(A). That is, the definition of the projection signs and the way the summation is performed in practice should be problem-specific and can be far from trivial, like in the case of ice crystals.

Unlike spherical harmonics which include several powers of both cosines and sines, vector correlations with directors are determined by projecting cosines only. This means that deviations from the selected set of orientations fade out rather slowly. In order to improve on the sensitivity of the projection procedure, in (8) we suggest collecting the maximum cosines

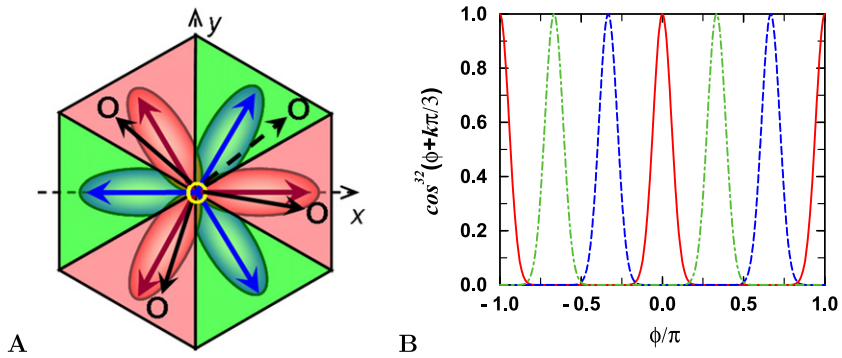

Figure 5. (A) Illustration of projecting onto a set of six directors within the xy-plane. The two incompatible director triplets are shown in different tones, the area of maximum correlation with each director being shaded accordingly. The xy-projection of four sample bonds around the central O-atom is given, where the dashed bond is to be damped out as inconsistent with the dominant pattern. (B) Function, $\cos^{32}(\phi + k\pi/3)$, $k = 0, 1, 2$, illustrates how the projections onto a set of directors in the xy-plane fade out in mid regions.

in some power, $n > 1$, where n is an adjustable parameter providing control over the desired strength of correlations. In the current case it is natural to require for projections to vanish in the mid region between the nearest-neighbour directors, i.e. about 30° away from each director. The latter condition is satisfied if one takes projecting cosines to the power of $n = 32$, see figure 5(B).

Let us now clarify how the summation in (8) can be done in practice so that the boat/chair confusion is avoided. The fact that all boat forms in ice I_h are aligned along the same axis (c-axis) gives us a nudge as to how to get around the issue. One way of dealing with it is to split the sum in (8) into two sums: one, $q_z^{(i)}$, for the projections of bonds that are roughly aligned with the z-axis and which are allowed to adopt either a boat- or a chair-like configuration (the directors with $z_k = \pm 1$), and the second sum, $q_{xy}^{(i)}$, for all the other projections that correspond to either of the two incompatible triplets of directors with $z_k = \pm 1/3$, taken with opposite signs. From here on we call these directors the *xy-triplets*. At this stage the sign of each contribution to either of the sums is set according to the normal antisymmetry rules, for xy-triplets antisymmetry being imposed within the projection onto xy-plane, as shown in figure 5(A). As soon as the counting is done for all neighbours of the central molecule, the *modulus* of the average correlation with the xy-triplets, $|q_{xy}^{(i)}|$, should be multiplied by the average z-sign of the corresponding OO-bonds and added with negative sign to the z-sum, $q_z^{(i)}$. This way the information about the dominant xy-triplet and the damped out bonds (inconsistent with each other) is preserved in the modulus of $q_{xy}^{(i)}$, whereas the correct z-antisymmetry is imposed by averaging over z-signs. Thus, we arrive at the following expression for the bond orientation parameter suited for both ice structures, I_h and I_c,

$$q_{\rm ice}^{(i)} = q_z^{(i)} - \langle \mathrm{sign}(z)\rangle_{xy}^{(i)} |q_{xy}^{(i)}|, \tag{9}$$

where the two sums read

$$q_z^{(i)} = \frac{1}{N_z^{(i)}} \sum_j \Pr_{\max}\{\mathbf{d}_{z=\pm 1}\}(\mathbf{u}_{ij}), \tag{10}$$

$$q_{xy}^{(i)} = \frac{1}{N_{xy}^{(i)}} \sum_j \mathrm{sign}_{xy}(\mathbf{d}_{z=\pm\frac{1}{3}}) \Pr_{\max}\{\mathbf{d}_{z=\pm\frac{1}{3}}\}(\mathbf{u}_{ij}), \tag{11}$$

here the maximum projection of $\mathbf{u}_{ij}$ onto the corresponding director set is

$$\Pr_{\max}\{\mathbf{d}_{(\cdots)}\}(\mathbf{u}_{ij}) = \sum_{k(\cdots)} \begin{cases} (\mathbf{u}_{ij}\mathbf{d}_k)^n & \text{if } \mathbf{u}_{ij} \text{ is max. corr. with } \mathbf{d}_k, \\ 0 & \text{otherwise,} \end{cases}$$

while the antisymmetry of xy-triplets is defined by

$$\mathrm{sign}_{xy}(\mathbf{d}_{z=\pm\frac{1}{3}}) = \begin{cases} \mathrm{sign}(x_{ij}) & \text{if max. corr. with } \mathbf{d}_{y=0,z=\pm\frac{1}{3}}, \\ -\mathrm{sign}(x_{ij}) & \text{if max. corr. with } \mathbf{d}_{y\neq 0,z=\pm\frac{1}{3}}, \end{cases}$$

and the average z-sign of $\mathbf{u}_{ij}$ vectors projected onto the xy-triplets is

$$\langle \mathrm{sign}(z)\rangle_{xy}^{(i)} = \frac{1}{N_{xy}^{(i)}} \times \sum_j \sum_k \begin{cases} \mathrm{sign}(z_{ij}) & \text{if max. corr. with } \mathbf{d}_{z=\pm\frac{1}{3}}, \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

$N_z^{(i)}$ and $N_{xy}^{(i)}$ being the accumulated numbers of projections onto the corresponding director sets.

Since the above formulae are rather involved to grasp at first glance, below we present the algorithmic equivalent thereof, which should be easier to follow. We notice that there are only three basis vectors which, when subjected to appropriate symmetry operations, are sufficient for producing the 14 directors that span over the vertices of all the tetrahedral configurations of interest. In terms of Euclidean coordinates all these symmetry operations can be traced to mere sign alternations in (7). Therefore, for a given (pre-normalized) vector between two O-atoms, $\mathbf{u}_{ij} = (x_{ij}, y_{ij}, z_{ij})$, it appears convenient to skip, at first, the signs of all its coordinates and find out only the maximum correlation with one of the three basis directors, $\{\mathbf{d}_k^{(0)}, k = 1, 2, 3\}$, (7) wherein all the coordinates are taken positive. This is done by finding k that gives $d_{\rm corr}(\mathbf{u}_{ij}) \equiv \max\{|x_{ij}d_{k,x}^{(0)}| + |y_{ij}d_{k,y}^{(0)}| + |z_{ij}d_{k,z}^{(0)}|\}$. The

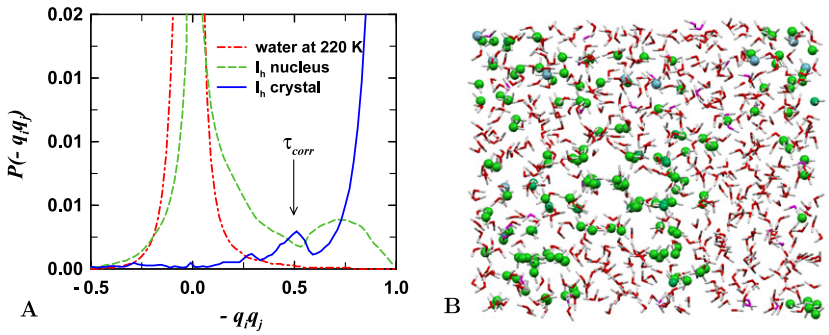

Figure 6. (A) Probability distributions of the product, $-q_{\rm ice}^{(i)}q_{\rm ice}^{(j)}$, in supercooled bulk water ((red) dot–dashed line), in a water system containing an ice I_h nucleus ((green) dashed) and in the I_h crystal ((blue) solid), all at 220K. (B) Illustration of an ice-like cluster (with O-atoms depicted as balls) found in supercooled bulk water by connecting all O-atoms for which $-q_{\rm ice}^{(i)}q_{\rm ice}^{(j)} > 0.5$ (the system containing 768 TIP4P molecules).

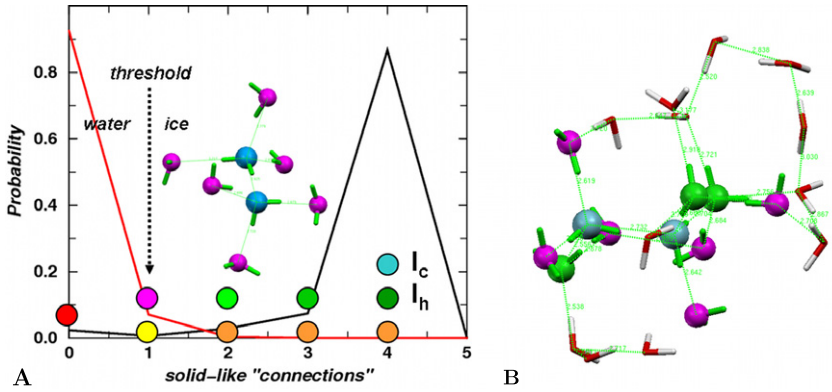

Figure 7. (A) Probability distributions for the number of ice-like connections in supercooled water (red line) and ice (black) at 220K. The colour code for molecules with different degree of connectivity is also shown: the upper row of circles is for O-atoms in the largest cluster (green tones and blue) and its interfacial particles (magenta), the lower row of circles for the rest of the system (red, yellow and orange). (B) A cut from a simulation snapshot illustrating a small ice-like cluster supported by insufficiently correlated water molecules.

latter quantity is, then, included into the corresponding sum in (9) but with an appropriate sign that identifies the actual director (rather than the basis director).

Thus, the scheme is: (i) for each bond identify the basis director ($k = 1$, 2 or 3) that gives the maximum projection, $d_{\rm corr}(\mathbf{u}_{ij})$; (ii) include this contribution into (9) with the aid of the following rules:

if $k = 1$, then $\mathrm{sign}(z_{ij})d_{\rm corr}^{n}(\mathbf{u}_{ij})$ is added into $q_z^{(i)}$, (10);
if $k = 2$, then $\mathrm{sign}(x_{ij})d_{\rm corr}^{n}(\mathbf{u}_{ij})$ is added into $q_{xy}^{(i)}$, (11);
if $k = 3$, then $-\mathrm{sign}(x_{ij})d_{\rm corr}^{n}(\mathbf{u}_{ij})$ is added into $q_{xy}^{(i)}$, (11).

(iii) In the latter two cases also add $\mathrm{sign}(z_{ij})$ to the sum in (12). Upon counting all the neighbours, the averages are collected in (9). The condition for the shells of two water molecules being correlated to produce an ice-genic connection is then very simple:

$$-q_{\rm ice}^{(i)}q_{\rm ice}^{(j)} > \tau_{\rm corr}. \qquad (13)$$

The effectiveness of the approach is illustrated in figure 6 which shows the probability distributions of the product, $-q_{\rm ice}^{(i)}q_{\rm ice}^{(j)}$, for supercooled bulk water, a large ice I_h cluster in supercooled water and the totally crystallized system, all at the same temperature $T = 220$ K. In the first case the distribution has its primary maximum around zero, corresponding to prevailing unstructured water, while the presence of the dissolved ice-like structure is revealed by the secondary maximum where $-q_{\rm ice}^{(i)}q_{\rm ice}^{(j)} > 0.5$. According to our preference, the distribution for the ice crystal is almost completely squeezed into the region above 0.5, having only minor features at lower values which may be attributed to the diffuseness of the structure at $T = 220$ K. Also, in figure 7 the distributions for the number of solid-like connections per water molecule (O-atom), $N_{\rm cn}$, are given for both the supercooled bulk water and equilibrated ice I_h systems. The downwards arrow points to the threshold value above which, $N_{\rm cn} > N_\tau = 1$, a molecule is included into a solid-like cluster. Due to the fact that $-q_{\rm ice}^{(i)}q_{\rm ice}^{(j)}$ separates well icy structures from disordered water and, thus, $N_{\rm cn}$ is also a very sensitive measure of ice-genic networking, molecules having only two ice-like connections can be definitely considered as clustering. This is particularly important because, apart from possible small threads, a separate hexagonal ring contains only such minimally interconnected molecules. The molecules with only one connection, if not making up a simple dimer, are linked to some cluster and represent the interface layer supporting the cluster within undercorrelated water environment.

Here it is worth noting that the introduced order parameter, $q_{\rm ice}^{(i)}$, is *not* rotationally invariant. That is, it is *specifically designed and suited* only for the two ambient ice forms, I_h and I_c, where molecular configurations of the boat and chair type can alter only along a predefined axis, corresponding to the c-axis of either of the two ice lattices. By definition, the presented summation procedure, (9)–(12), is not supposed to effectively distinguish the structures of ice phases where the axis of the boat–chair dualism is not fixed (e.g. ice II) or other additional complications arise (e.g. due to the presence of pentagonal five-membered rings in ice III). It should not be difficult to adapt the procedure to those cases too. For example, one can gain the invariance with respect to the orientation of the duality axis by (re-)aligning the *local* z-axis with the OO-connection *in focus*, see figures 3(A) and 4—in this case no angular diffuseness around the z-axis should be assumed. We note, however, that for the purposes of the present study we do not need to make our orientational order parameter more flexible than it is, because, according to the TIP4P model phase diagram [1], it is rather unplausible to expect any other ice phase to interfere with the I_h and I_c forms at ambient pressure.

We now address how one discriminates between molecules belonging to a particular ice phase. As is seen in figure 4, the distinction can be based on the (anti)symmetry relations between the two tetrahedra forming a connection. That is, the two xy-triplets participating in the formation should obey the following rules: (i) they should be either symmetric (boat form) or antisymmetric (chair form) with respect to each other, and (ii) their average z-projections should have opposite signs. These criteria can be used in order to count the number of boat- and chair-like connections each molecule makes with its neighbours. Provided the primary condition for an ice-like connection, (13), is fullfilled, i.e. a given pair of molecules, i and j, is certainly involved in one of the two sought arrangements, we can decide on the connection type by using solely the first criterion for the xy-triplets,

$$q_{xy}^{(i)} q_{xy}^{(j)} \begin{cases} > 0 & \text{symmetric, boat form,} \\ < 0 & \text{antisymmetric, chair form.} \end{cases} \tag{14}$$

Now, knowing the number of boat and chair forms around each molecule, we have the means to identify its phase. Since boat configurations are completely absent in cubic ice and constitute only a quarter of connections in hexagonal ice, it is tempting to label all molecules that are found in at least one boat-like arrangement as I_h phase, and I_c otherwise. It is clear, though, that initially loose (liquid-like) molecules are likely to pass through intermediate unresolved states, and the type of each connection may change a few times before a molecule settles down within a larger crystal structure. So it appears meaningless trying to label a water molecule unless it has at least three sound ice-like connections with its neighbours. We chose to apply an even stricter rule: only the molecules with all four connections (and hence H-bonds) established are classified. In all our simulation snapshots molecules in the I_c environment are depicted in blue, and those in I_h structure are dark-green, see figure 7 for the full colour code with respect to the number of ice-genic connections.

To summarize, the advantages of the maximum projection method are: (i) it is relatively simple and straightforward to code; (ii) it is rather general and can be adapted to other orientation analysis problems (one has to merely redefine the directors and the sign rules in (8)); (iii) it is particularly suited for low-symmetry cases where the spherical harmonics analysis may require inclusion of high order harmonics or may fail completely; (iv) the director scheme is easily understood in terms of ordinary 3D vectors.

4. Monte Carlo simulation: competitive growth of ices I_h and I_c

4.1. Choosing the coordinate for bias

By using the orientational order parameters introduced above, it is possible to find all clusters in the simulated system in the same manner as it is done for spherical particles. A general approach to simulating nucleation is to direct (bias) growth of the largest cluster, as opposed to global crystallization where biasing is applied to the entire system. That is, by focusing the bias on the largest available nucleus and ignoring all the smaller ones (allowing those to dissolve), we, presumably, reduce the probability for two nuclei to grow near each other and, therefore, we can hope that any interaction between them does not affect the resulting free energy barrier [20, 21]. In the case of ice it proved, however, inefficient to use the number of molecules in the largest cluster, $N_{\rm nuc}$, as the coordinate for biasing. This can be understood with the aid of figure 6, from which it is clear that transforming a molecule between liquid-like and crystal-like states is associated with gradually increasing the corresponding $-q_{\rm ice}^{(i)} q_{\rm ice}^{(j)}$ products. This process appears to be much slower and more tricky in water environment than in simple fluids, amounting essentially to stepwise accumulation of 'good enough' ice-like connections. As a side issue, there is always a broad distribution of the ice-likeness within a cluster, which means that two clusters with the same number of water molecules can differ considerably in the quality of their structures. Obviously, this leads to ambiguities and inevitable hysteresis in both the biasing procedure and the estimated free energy. It is to say that, unfortunately, increasing the threshold in (13) does not help much because this way one only makes it harder for a molecule to be included into a cluster, whereas the aim is to smoothly grow the latter. In order to minimize such effects, we introduce a quality criterion and replace $N_{\rm nuc}$ with an *effective* nucleus size, $N_{\rm eff}^{(\rm nuc)}$, defined by

$$q_{\rm eff}^{(\rm nuc)} \equiv \frac{N_{\rm eff}^{(\rm nuc)}}{N} \equiv \frac{1}{N} \sum_{i}^{N_{\rm nuc}} \frac{1}{N_{\rm b}^{(i)}} \sum_{j}^{N_{\rm b}^{(i)}} (-q_{\rm ice}^{(i)} q_{\rm ice}^{(j)}), \tag{15}$$

where we have also introduced the *reduced* effective size of the nucleus, $q_{\rm eff}^{(\rm nuc)}$; as above, $N_{\rm b}^{(i)}$ is the number of bonds (not connections!) for molecule i within the nucleus. One can notice that, when applied to the entire system, $q_{\rm eff}^{(\rm nuc)}$ turns into a global order parameter spanning values between -1 and 1.

4.2. Simulation details

All simulations presented below were carried out in an NpT ensemble with $p = 0.1$ MPa, and the TIP4P water model with Ewald summation for electrostatic interactions was employed, which is not only the most popular water model but also seems to reproduce well the shape of the experimental phase diagram of water (up to some scaling) [1]. As in similar studies reported earlier by other groups [18, 19], the geometry of the periodic simulation box was chosen commensurate with the sought ice crystal(s). Isotropic volume changes, retaining ratios between the box dimensions, were attempted randomly in $\ln(V)$ at a rate of $1/N$, i.e. one volume step per MC sweep (on average). While in a sufficiently large system the barrier to homogeneous nucleation should not be affected by the shape of the simulation box, at the current stage the problem of ice crystallization/nucleation remains extremely computationally demanding, which puts severe restrictions on the system size that can be simulated. Thus, maintaining the system in a shape matching the prospective ice lattice(s) is intended for minimizing any unphysical boundary effects which might arise when simulating a system restricted in size. No doubt, one has to be careful in interpretation of the data obtained in such circumstances and, therefore, in this report we do not expect to attain the actual nucleation barrier. Our primary goal here is rather to establish an efficient and reliable simulation method for directing homogeneous ice nucleation in the least biased manner.

Up to now systems consisting of $N = 96$, 360 and 768 water molecules with periodic box dimensions matching the (stable) ice I_h structure have been simulated below the temperature of melting, $T_{melt} = 230$ K [1, 25]. In some cases, especially with the largest system size, for the sake of considerable increase in simulation speed (up to 5–6 times) we skipped the reciprocal part in the Ewald sum. Nowadays it has been established that in dense ionic mixtures—water being a typical example—due to dramatic screening of electrostatics, local charge distributions can be well approximated by renormalized (reduced) charges so that partially charged molecules in solution can be seen as essentially neutral objects when examined from sufficiently long distances [43]. Particularly, in the case of water the long-range electrostatic term contributes only a tiny percentage into the total energy of a molecule and skipping it does not affect the TIP4P model properties notably[4].

In order to obtain the Gibbs free energy barrier(s) with respect to $N_{eff}^{(nuc)}$, or equivalently $q_{eff}^{(nuc)}$, MC simulations of umbrella sampling type were performed in a few sets of parallel runs, each covering a restricted sub-range of $q_{eff}^{(nuc)}$. Each parallel simulation was carried out in several overlapping windows (smaller sub-ranges), with a *unified* harmonic bias potential being applied with respect to $q_{eff}^{(nuc)}$. A replica-exchange (RE-MC) technique was used to allow for movements between adjacent windows [20, 21]. By a unified bias potential we mean that it is a *single* parabola over the entire range of $q_{eff}^{(nuc)}$, which is in contrast to several parabolic wells of different width, each centred in the mid-point of its own window, being most often used in simulations of this kind. The advantages and essential details of using a unified umbrella potential are discussed in the following section.

4.3. Replica-exchange umbrella sampling with unified bias potential

In general, the purpose of using a bias potential is to effectively compensate for the underlying free energy (sometimes entropy) landscape so that the system could explore more freely the configuration space and, thereby, overcome barriers and reach otherwise rarely visited (sub-)states. As such, the sought bias, or penalty function, has always been a sort of the Holy Grail in the domain of self-adjusting free energy methods [28–32]. In contrast, with umbrella sampling [26, 27] it became a rule of thumb not to seek the elusive ideal but routinely apply strong enough parabolic umbrellas in sufficiently narrow windows [20–22], which works robustly, although it is not the optimum approach on the slope of a steep barrier. That is, in the latter case only half of the parabola actually drives the system onto the barrier whilst the other half in effect makes the sensed slope even steeper! Clearly, as the barrier slope steepens one has to progressively squeeze the windows and ensure that they overlap, or at any rate touch, by the favourable parts of the corresponding umbrellas. This would not be the case if one knew (or could guess) at least the direction of the free energy rise, because then one could put the centre of each harmonic well not in the window mid-point but at one of the edges (or even further out) where the free energy is assumed to be higher. This way the bias would be favourable in a whole window. A natural extension of this 'stitched umbrellas' approach is to use a single harmonic well over the range where the free energy is thought to be increasing monotonously and steeply. Again, ideally the centre of the well is to be placed as close to the actual barrier top as possible. In our case, since the system sizes are (most likely) too small for actually observing the critical nucleus, we put the umbrella centre close to $q_{eff}^{(nuc)} = 0.5$, unless otherwise specified.

A nice 'drawback' of using a unified umbrella in conjunction with replica-exchange is that there is no change in the model Hamiltonian associated with an exchange of configurations between overlapping windows. This is because no hopping between (different) parabolas actually occurs. The exchanged replicas merely swap their $q_{eff}^{(nuc)}$-ranges, or window identities. Of course, for such a swap move to be possible a pair of configurations must be found within the overlapping region of two adjacent windows.

Assuming that each window has been explored by the system sufficiently well, i.e. every $q_{eff}^{(nuc)}$-bin has been visited at least a few times (say, $N_{vis} > 10$), the collected histogram of visits, $N_{vis}(q_{eff}^{(nuc)})$, allows for calculating the relative free

[4] We expended some effort to study the effects of skipping the reciprocal Ewald sum on the energy, structural and thermodynamic properties of the system. We found that by carefully selecting the real space cut-off ($R_{cut} = 8.5$ Å) and the screening parameter ($\alpha = 0.3$) it was possible to keep the energy error below 0.5%, whilst no deviation in the RDFs was observed as compared to the case with full Ewald sum. Independent MD simulations of close to coexistence water and ice I_h phases did not reveal any drift in T_{melt} either.

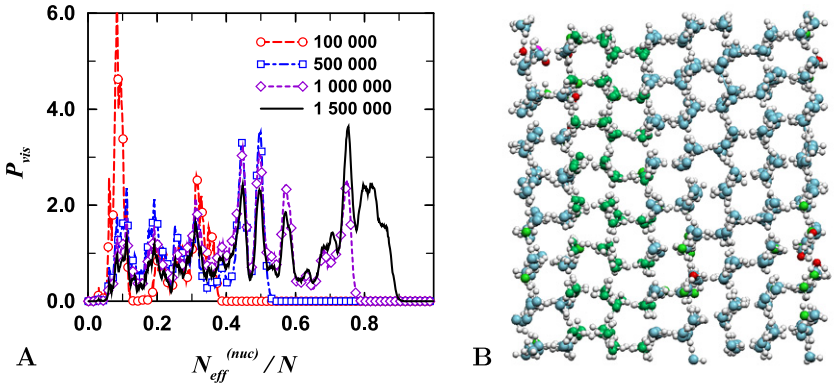

Figure 8. (A) Sampling probability distributions for the system of 768 molecules during the process of a biased ice nucleation. Subtotal distributions progressively collected at different stages are shown. The legend displays the number of MC steps per molecule. (B) The final mixed ice structure.

energy within kth window as

$$\beta G_k\left(q_{\rm eff}^{\rm (nuc)}\right) = -\beta W_k^{\rm (bias)}\left(q_{\rm eff}^{\rm (nuc)}\right) - \ln\frac{N_k^{\rm (vis)}(q_{\rm eff}^{\rm (nuc)})}{N_k^{\rm (tot)}} + {\rm const}, \tag{16}$$

where $\beta = 1/(kT)$, $W_k^{\rm (bias)}(q_{\rm eff}^{\rm (nuc)})$ is the applied bias potential (in our case the same harmonic function), and $N_k^{\rm (tot)}$ is the total number of samples in the window. The differences between unknown constants can always be eliminated from the consideration by use of an appropriate stitching procedure for the continuous pieces of the free energy function obtained within adjacent windows. Indeed, one can notice that the possible discontinuities in $G(q_{\rm eff}^{\rm (nuc)})$ can develop only at the edges of windows, which are in our case separated by half-a-window intervals in $q_{\rm eff}^{\rm (nuc)}$. Thus, in order to (re-)construct the continuous free energy profile in the entire parameter range, one has to suitably shift the $G(q_{\rm eff}^{\rm (nuc)})$ sub-ranges.

Below we use index k not for windows but for half-a-window intervals, whereas the order parameter dependence is replaced by a discretized notation with a superscript index for $q_{\rm eff}^{\rm (nuc)}$-bins. Then, starting with $k = 2$ and subsequently incrementing k, we chose to update all the free energy values obtained within kth interval by adding

$$\Delta G_{k-1,k}^{\rm (edge)} = \tfrac{1}{2}\left[\left(G_{k-1}^{(m)} - G_{k-1}^{(m-1)}\right) + \left(G_k^{(2)} - G_k^{(1)}\right)\right] - \left(G_k^{(1)} - G_{k-1}^{(m)}\right), \tag{17}$$

where m is the maximum bin index in each interval. Obviously, this procedure corresponds to a (mean) linear extrapolation for the free energy difference between the adjacent utmost bins in kth and $(k-1)$th intervals. Although more sophisticated schemes can be applied [27], we found that (17) produces quite a satisfactory continuous $G(q_{\rm eff}^{\rm (nuc)})$-profile, without detectable abrupt inflections.

4.4. Biased ice crystal growth

Prior to free energy calculations, we had to obtain initial configurations corresponding to the values of $q_{\rm eff}^{\rm (nuc)}$ falling within each window so as to cover the entire range between the states of liquid water and ice. This task by itself appeared rather tedious in terms of the simulation effort. For each system size we performed several simulations starting with equilibrated supercooled water configurations and using different random sequences. Not to overdamp the dynamics of sampling, we chose for this stage a rather modest supercooling regime, $T = 220$ K—just below $T_{\rm melt} = 230$ K. A very steep biasing potential had to be applied in order to drive the system over the crystallization barrier. In these simulations we put the parabola centre at $q_{\rm eff}^{\rm (nuc)} = 0.8$ whereas its height at the liquid end reached hundreds kTs (1200 kT for $N = 768$). While with 96 molecules about 10^5 MC sweeps (steps per molecule) were necessary to attain total crystallization of the system, the required simulation length increased progressively with the system size: it tripled for 360 molecules, and over 10^6 steps per molecule were needed for ice structures to build up with $N = 768$. The sluggish sampling of $q_{\rm eff}^{\rm (nuc)}$ during this long-term process is illustrated for the latter system in figure 8.

In all cases we eventually obtained pure hexagonal ice and mixed structures with alternating sheets of I_h and I_c ices in different proportions, as is illustrated in figure 9 for the system of 360 molecules. In neither case did the pure cubic ice arise. This is not surprising with $N = 96$ and 768, owing to the fact that in these systems the periodic box does not fit I_c lattice along the z-axis. However, with $N = 360$ the box dimensions actually match the continuous cubic ice lattice. Yet, it was not obtained.

In figure 10 a few clusters found at initial stages of nucleation with $N = 768$ are presented, see also figures 6(B) and 7(B). First of all, we clearly see that the maximum projection procedure presented in section 3 works well. Second, the picture suggests that, despite the method of biasing the largest cluster growth being used, already at initial stages of crystallization many small ice-like arranged threads and rings are found, which later merge into a few separate but apparently correlated clusters. Evidently, these long-range correlations must be imposed by H-bond networks which, eventually, spread over the entire system and settle down into the overall crystal structure. Note how strong short-range correlations [44] produce a long-range network of clusters which is gradually self-tuning into a large-scale ice lattice. At some intermediate stage of this 'tuning-up'

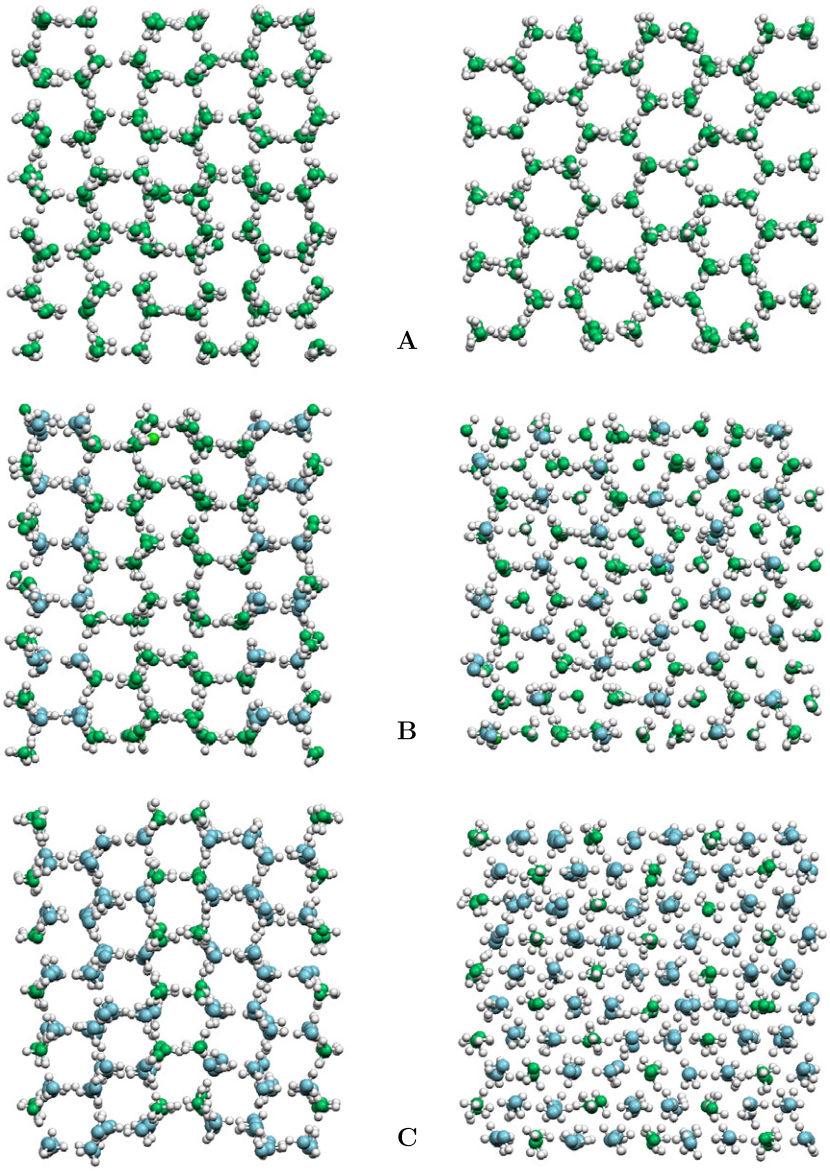

Figure 9. Three ice structures obtained with 360 (TIP4P) water molecules at 220 K. (A) hexagonal I_h phase; (B) and (C) mixed phases with different proportions of alternating hexagonal and cubic structures, I_h:I_c = 2:1 and 1:2, respectively. The colour code: green for I_h and blue for I_c. The prism face is shown in the left column, and the basal face in the right.

process a well-structured crystal core emerges, similar to that in figure 6(B). It is worth mentioning that when looking at snapshots where all water molecules are present and only the largest cluster is highlighted, it is easy to miss out the larger surrounding network because smaller clusters are concealed by undercorrelated water molecules.

These observations, whilst being generally in line with the arguments given by Matsumoto *et al* [15], seem to be, in some respect, at odds with the findings of Radhakrishnan and Trout for biased crystallization of pure hexagonal ice (JACS) [18] as well as the metadynamics results of Quigley and Rodger for pure cubic ice [19]. In these studies similar multidimensional global order sets (2D and 3D, respectively, both included Q_6 and some tetrahedral correlation measure) were used for biasing and yielded different ice forms. Furthermore, the authors of both papers concluded that they observed crystallization proceeding with a compact well-shaped nucleus. However, in the first case, [18] where a system of about a thousand molecules was treated, it is difficult to judge the initial stages of nucleation, as no detailed analysis of the nucleus size and geometry was presented[5] and only simulation snapshots of the entire system were given. By inspecting the snapshots (figure 1 in their JACS paper) one can see that about *half way* on the barrier slope (figure 2) the system adopts configurations where some hexagonal patterns can be detected by naked eye (figure 1b). These are visible because a larger-scale arrangement—in the shape of an empty tube—has, in fact, spread across the simulation box, as Quigley and Rodger also noticed [19]. In the next snapshot (figure 1c), corresponding to a small increment in Q_6, the ice structure

[5] In this study only the order parameter distributions for local 'clusters' consisting of nearest-neighbour molecules were given.

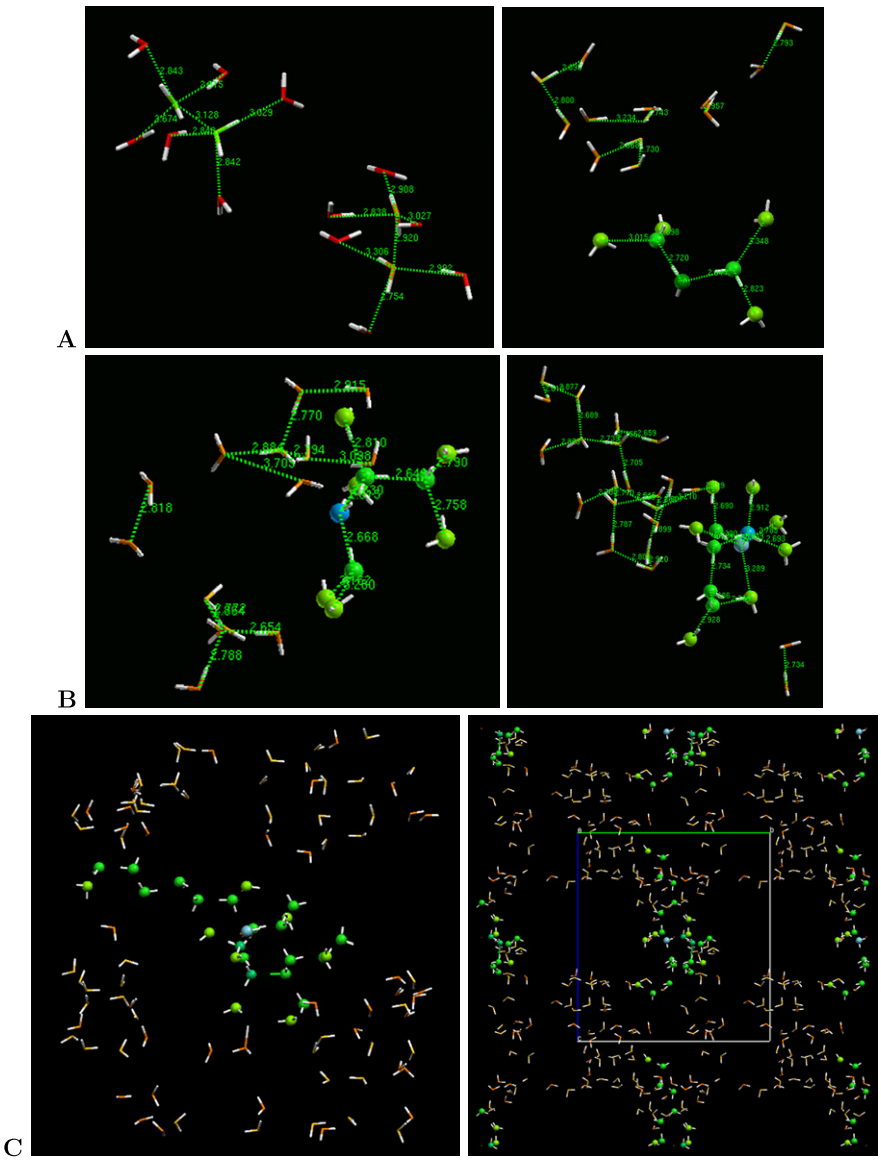

Figure 10. Snapshot cuts of ice clusters found in the system of 768 (TIP4P) water molecules at different stages of nucleation. (A) small initial structures; (B) tetrahedrally arranged threads and rings; (C) cluster network formation (the same configuration viewed at different angles and scale). Insufficiently correlated water molecules have been removed for clarity.

has already blown and become well pronounced, and one can anticipate its percolation through the volume in all directions, similar to what is seen in the snapshot we present in figure 6(B). A typical configuration shown by Radhakrishnan and Trout for the top of the barrier (figure 1e) is even more structured and spread—the periodic boundary effects must be in play here. Thus, the overall crystallization process observed by these authors is not, actually, at variance with the image of long-range correlations spreading over the network of ice-like structures, as is found by use of the cluster analysis in all our systems irrespective of the obtained ice lattice. In the work of Quigley and Rodger [19] the system size was rather small (576 molecules), which makes it even more difficult to draw a definite conclusion about the observed freezing process. Their snapshots reveal, though, that in their case too, after a long initiation period with small unshaped clusters present, the ice structure eventually propagated in all directions rather easily. Note that, despite very similar sets of order parameters being used, in neither of the two studies combined structures of ices I_h and I_c were obtained, whereas we constantly observe mixing of the two ice forms in our simulations. We therefore infer that the current mosaic picture of ice nucleation extractable from the literature is far from being complete or satisfactory. It is also clear that larger systems need to be simulated in order to gain conclusive data on the formation of the critical nucleus, its size and shape.

4.5. Free energy calculations

Starting with the obtained initial configurations, we performed equilibration runs with 10^5 MC steps per molecule in every window. These were done in essentially the same fashion as the

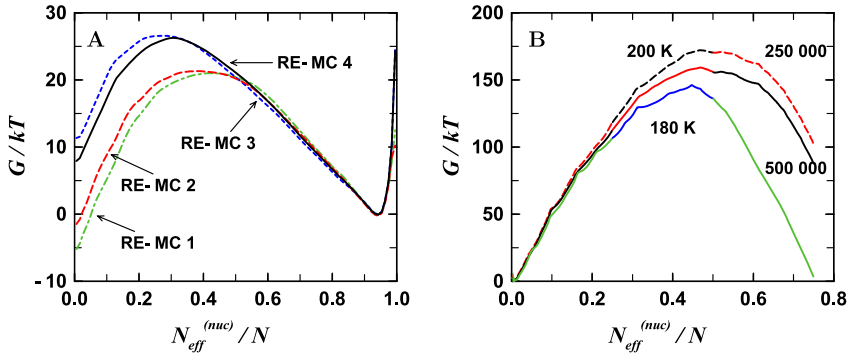

Figure 11. Gibbs free energy barriers to ice formation obtained for: (A) the system with $N = 96$ at $T = 220$ K (each line represents a separate simulation of over 1 million MC sweeps); (B) the system with $N = 768$ at $T = 200$ and 180 K after 0.5 million MC sweeps (dashed are the intermediate results, $T = 200$ K).

prospective free energy calculations, by use of parallel replica-exchange MC simulations with a moderate harmonic biasing potential centred at $q_{\rm eff}^{(\rm nuc)} = 0.5$. The height of the potential at $q_{\rm eff}^{(\rm nuc)} = 0$ was 175 kTs for $T = 200$ K and 150 kTs for $T = 180$ K. In all cases the discretization of the order parameter was such that each window consisted of 16 bins.

The current free energy data have been acquired mostly for $N = 96$ and the large system, $N = 768$, that crystallized into a mixed ice structure shown in figure 8(B). The resulting Gibbs free energy profiles, $G(q_{\rm eff}^{(\rm nuc)})$, are given in figure 11. Since the free energy is evaluated only up to a constant, the relative positioning of its profiles is essentially a matter of a reasonable convention. Thus, the choice of zero levels in the two plots was guided by different considerations, as explained below. Again, judging by the common shape of all the lines we conclude that biasing with respect to the effective cluster size does bring the system over the nucleation barrier.

Strikingly, with the smallest system size we observe dramatic variations in the barriers obtained in subsequent simulations, each over a million steps per molecule (in each window). This is in contrast to our naive expectation that due to the imposed artificial periodicity the smaller system should have behaved nicer, displaying more stable statistics than the larger ones. Such a diverse spread in the barrier shape and height might lead one into thinking that the final ice structure must have changed its identity from run to run. It was not the case, however, which is revealed by the well-shaped minimum corresponding to ice, $q_{\rm eff}^{(\rm nuc)} > 0.5$, where only slight deviations of the $G(q_{\rm eff}^{(\rm nuc)})$ slope are seen. Indeed, on this side of the barrier a well-developed ice structure is constantly present and configurational fluctuations are strongly damped in comparison to those observed at the other side. To emphasize this observation, in figure 11(A) we put all the ice phase minima at zero level so that the variations in $G(q_{\rm eff}^{(\rm nuc)})$ span over the region in $q_{\rm eff}^{(\rm nuc)}$ where they originate from.

For the larger system, $N = 768$, we do not have as much statistics and the obtained free energy data do not cover the entire $q_{\rm eff}^{(\rm nuc)}$-range (which is not necessary for estimation of the barrier). In figure 11(B) the $G(q_{\rm eff}^{(\rm nuc)})$-profiles are placed in a more conventional way, i.e. their minima for the liquid (parent) phase are fixed at zero. It is clear that in this case the estimated barriers are unlikely to be settled after only half-a-million MC sweeps. The difference between the dashed and solid lines, which correspond to the partial data obtained at $T = 200$ K after 0.25 and 0.5 million sweeps, shows an obvious trend of reduction in the barrier height. This trend indicates that the system may not have been equilibrated yet, which is then even more true at $T = 180$ K. In the latter case some wavy kinks on the barrier slope disclose the presence of long-lived structures in supercooled water—a typical picture observed in partial data for all system sizes.

Regarding the location of the barrier tops in figure 11(B), which roughly correspond to half the system size, $q_{\rm eff}^{(\rm nuc)} \approx 0.5$ (the actual $N_{\rm nuc}/N$ being somewhat larger), we inevitably conclude that the simulated system is way too small to embrace the critical nucleus. This is in accord with the afore-discussed cooperative tuning of small clusters which progressively merge into/with larger ones—the process that in a small system eventually results in a relatively fast spread of ice-genic H-bonding on the global scale. Thus, as in the earlier studies [18, 19], a (sudden) percolation of the coherent ice structure through the periodic boundaries is normally seen in configurations corresponding to the transition state. In the light of this fact, we have to stress that all the fee energy data for ice crystallization reported up to date were obtained for undersized systems.

Even though it is evident that our free energy data are insufficient to be conclusive, our rough estimate for the height of the nucleation barrier for the largest simulated system, falling within 120–150 kT, is way too large in comparison to the barrier heights reported earlier for systems of comparable sizes: 60 kT for hexagonal ice [18] and 70 kT for cubic ice [19]. While we cannot with certainty explain the discrepancy, we realize that one cause may be the difference in the biasing methods. In both the referred works similar sets of *global* order parameters were employed with the implication that the estimated *crystallization* barriers may be actually not related to the true *nucleation* pathway. Apart from that, in those simulations only pure ice phases, either hexagonal or cubic, were emerging, whereas in our study we focused on growing the largest cluster and (successfully) enabled both ice forms to arise, without favouring a particular phase. The latter aspect seems to be important because allowing the two ice structures

to combine is likely to encourage a greater variety of ice-like arrangements. Then, the initiated clustering process may temporarily get stuck, thereby making the free energy barrier higher. Moreover, the issue of incoherent growth of the two ice forms should be especially pronounced in small systems, as soon as the developing structure starts percolating through the periodic boundaries—the effect that is certainly present with the current system sizes, e.g. see figure 6(B). We note also that in molecular dynamics simulations of the brute-force type [17] the interface between two separate—and, thus, incoherent—nuclei persists in excess of 0.5 μs. In contrast, if a single ice form is favoured, it is plausible to expect its propagation through the volume going easier, because in this case molecular patterns at the opposite sides of the same cluster are most likely to be matching and, thus, readily stitching due to the box periodicity.

5. Concluding remarks

In summary, we have tackled the problem of ice formation in pure bulk water for which no unified bond orientation order parameter was available previously. As an alternative to spherical harmonics analysis, we introduced a maximum projection method and successfully applied it in the simulation of competitive growth of hexagonal and cubic ice forms. The suggested approach is sufficiently general and especially suited for cases where supramolecular arrangements can comprise patterns of different symmetries for which the conventional search for site-to-site orientation similarities may be error-prone.

Concerning our simulations of ice nucleation and calculation of the associated free energy barriers, altogether our observations are in accord with previous simulation studies of the same phenomenon. In particular, we observe how strong short-range correlations within relatively long-lived hydrogen-bonded networks [44] give rise to severe viscosity of supercooled water and depress the dynamics of sampling in simulation. We believe that for the same reason non-distant local crystallization centres in water appear to be correlated and, as such, gradually merge into a large-scale unshaped ice-genic network. Eventually, after H-bonds within this network sufficiently self-tune, a well-shaped ice core arises, being surrounded by a relatively structured and spread buffer zone which makes further growth easier. This description in large part follows the discussion of spontaneous ice I_h nucleation of Matsumoto *et al* [15] and suggests that the process is initiated by a collective fluctuation in H-bonding arrangement within a relatively large volume. Unfortunately, such a highly cooperative underlying mechanism results in a surprisingly effective percolation of the developing ice patterns through the simulation box, as is clearly seen in the transition state configurations for all system sizes studied up to now both in this report and by other groups [18, 19]. This is indicative of unsolved challenges in simulation of ice nucleation. Thus, a conclusive study of both the nucleation pathways and barriers would require longer simulations ($\sim 10^8$ steps/mol.) for larger systems ($\sim 10^4$ mol) than have been done currently.

Acknowledgments

This work has been supported by the EPRS grant GR/T27112/01. AB expresses his gratitude for helpful discussions of various aspects of this work to Dr Magnus Ullner (Lund University, Sweden), Professor Roland Kjellander (Göteborg University, Sweden) and Professor Pavel Jungwirth (Inst. of Org. Chem. and Biochem., Czech Republic).

References

[1] Sanz E, Vega C, Abascal J L F and MacDowell L G 2004 *Phys. Rev. Lett.* **92** 255701
Vega C, Sanz E and Abascal J L F 2005 *J. Chem. Phys.* **122** 114507
[2] Buch V, Martonak R and Parrinello M 2005 *J. Chem. Phys.* **123** 051108
Buch V, Martonak R and Parrinello M 2005 *J. Chem. Phys.* **124** 204705
[3] Petit J R, Duval P and Lorius C 1987 *Nature* **326** 62
[4] Schenk P, Matsuyama I and Nimmo F 2008 *Nature* **453** 368
[5] Sastry S 2005 *Nature* **438** 746
Shaw R A, Durant A J and Mi Y 2005 *J. Phys. Chem.* B **109** 9865
[6] Sassen K 2005 *Nature* **434** 456
Kaufman Y J, Tanré D and Boucher O 2002 *Nature* **419** 215
Murray B J, Knopf D A and Bertram A K 2002 *Nature* **434** 202
[7] Ascer J P, Elliott J A W and MacGann L E 2001 *J. Biophys.* **81** 1389
Blow N 2007 *Nature* **448** 959
[8] Campbell D B, Campbell B A, Carter L M, Margot J-L and Stacy N J S 2006 *Nature* **443** 835
[9] Bandfield J L 2007 *Nature* **447** 64
Schorghofer N 2007 *Nature* **449** 192
[10] Ostrovskii V E and Kadyshevich E A 2007 *Phys. Usp.* **50** 175
Ostrovskii V E and Kadyshevich E A 2007 *Adv. Phys. Sci. Rus.* **117** 183
[11] Tabazadeh A, Djikaev Y S and Reiss H 2002 *Proc. Natl Acad. Sci.* **99** 15873
[12] Meyer J C, Girit C O, Crommie M F and Zettl A 2008 *Nature* **454** 319
[13] Koga K, Tanaka H and Zeng X C 2000 *Nature* **408** 564
[14] Wales D J, Doye J P K, Miller M A, Mortenson P N and Walsh T R 2000 *Adv. Chem. Phys.* **115** 1
[15] Matsumoto M, Saito S and Ohmine I 2002 *Nature* **416** 409
[16] Yamada M, Mossa S, Stanley H E and Sciortino F 2002 *Phys. Rev. Lett.* **88** 195701
[17] Vrbka L and Jungwirth P 2006 *J. Phys. B: At. Mol. Opt. Phys.* **110** 18126
Vrbka L and Jungwirth P 2007 *J. Mol. Liq.* **134** 64
Bauerecker S, Ulbig P, Buch V, Vrbka L and Jungwirth P 2008 *J. Phys. Chem.* C **112** 7631
[18] Radhakrishnan R and Trout B L 2003 *J. Am. Chem. Soc.* **125** 7743
Radhakrishnan R and Trout B L 2003 *Phys. Rev. Lett.* **90** 158301
[19] Quigley D and Rodger P M 2008 *J. Chem. Phys.* **128** 154518
[20] ten Wolde P R, Ruiz-Montero M J and Frenkel D 1996 *J. Chem. Phys.* **104** 9932
ten Wolde P R and Frenkel D 1998 *J. Chem. Phys.* **109** 9901
[21] Auer S and Frenkel D 2001 *Nature* **409** 1020
[22] Chopra M, Muller M and de Pablo J J 2006 *J. Chem. Phys.* **124** 134102
[23] Desgranges C and Delhommelle J 2006 *J. Am. Chem. Soc.* **128** 10368
Desgranges C and Delhommelle J 2007 *J. Am. Chem. Soc.* **129** 7012
[24] Sanz E, Valeriani C, Frenkel D and Dijkstra M 2007 *Phys. Rev. Lett.* **99** 055501

Mladek B M, Charbonneau P and Frenkel D 2007 *Phys. Rev. Lett.* **99** 235702
[25] Koyama Y, Tanaka H, Gao G and Zeng X C 2004 *J. Chem. Phys.* **121** 7926
Wang J, Yoo S, Bai J, Morris J R and Zeng X C 2005 *J. Chem. Phys.* **123** 036101
[26] Torrie G M and Valleau J P 1977 *J. Comput. Phys.* **23** 187
[27] Frenkel D and Smit B 1996 *Understanding Molecular Simulation* (San Diego, CA: Academic)
[28] Iba Y 2001 *Int. J. Mod. Phys.* C **12** 623
[29] Wang F and Landau D P 2001 *Phys. Rev.* E **64** 056101
[30] Laio A and Parrinello M 2002 *Proc. Natl Acad. Sci.* **99** 12562
[31] Shell M S, Debenedetti P G and Panagiotopoulos A Z 2003 *J. Chem. Phys.* **119** 9406
[32] Broukhno A V (Brukhno) 2003 Free energy and surface forces in polymer systems *PhD Thesis* Lund University, Sweden
[33] Brukhno A, Anwar J, Davidchack R and Handel R 2007 *Long Time Scale Dynamics Workshop (Loughborough, April 2007)* http://www.functional-coatings.org/LongTimeScaleTalks/Brukhno.pdf
[34] Steinhardt P J, Nelson D R and Ronchetti M 1983 *Phys. Rev.* B **28** 784
[35] Chau P L and Hardwick A J 1998 *J. Mol. Phys.* **93** 511
[36] Takahashi T 1982 *J. Cryst. Growth* **59** 441
[37] Fennell C J and Gezelter J D 2005 *J. Chem. Theory Comput.* **1** 662
[38] Cacciuto A and Frenkel D 2005 *J. Phys. Chem.* **109** 6587
[39] Starr F W, Angell C A and Stanley H E 2003 *Physica* A **323** 51
[40] Johari G P 2005 *J. Chem. Phys.* **122** 194504
[41] Zhang W X, He C, Lian J S and Jiang Q 2006 *Chem. Phys. Lett.* **421** 251
[42] Zheng L, Luo S N and Thompson D L 2006 *J. Chem. Phys.* **124** 154504
[43] Kjellander R and Ramirez R 2005 *J. Phys.: Condens. Matter* **17** S3409
Fennell C J and Gezelter J D 2006 *J. Chem. Phys.* **124** 234104
Denesyuk N A and Weeks J D 2008 *J. Chem. Phys.* **128** 124109
[44] Raiteri P, Laio A and Parrinello M 2004 *Phys. Rev. Lett.* **93** 087801