Connectivity of hydrogen bonds in liquid water

Robin L. Blumberg and H. Eugene Stanley

Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215

Alfons Geiger and Peter Mausbach

Institut für Physikalische Chemie, Rheinisch-Westfälische Technische Hochschule Aachen, D 5100 Aachen, West Germany

We report on a molecular dynamics (MD) study of the connectivity of hydrogen bond networks in liquid water, focusing primarily on the microscopic distribution functions giving the weight fraction of molecules belonging to a "net" of M molecules ($M = 1,2,3,...$). The MD data compare favorably—using no adjustable parameters—with predictions of random bond percolation theory. We also study the connectivity of those molecules with four intact hydrogen bonds, and compare the corresponding distribution functions with correlated-site percolation theory. Our analysis supports the proposal that when looking at the *bond* connectivity, water appears as a macroscopic space-filling network—as expected from continuum models of water. When looking at the correlated *site* percolation problem defined by the four-bonded molecules, water appears as a myriad of tiny ramified low-density patches, somewhat reminiscent of mixture theories and cluster models. In Appendix A, we find a strong correlation between the number of molecules within a sphere of radius r_c around a given molecule and the total interaction energy of that molecule with its neighbors residing in that sphere; for most choices of r_c, the energy becomes less negative when more molecules are in the sphere, in contrast to the behavior of a normal fluid. This result supports the finding of Geiger and Stanley that regions of *high* bond connectivity are correlated with regions of *low* density. In Appendix B we describe in detail how we adapt conventional percolation theory to the calculation of cluster size distribution functions for hydrogen bond networks in water.

I. INTRODUCTION

It has long been recognized that the extensive degree of hydrogen bonding in liquid water is related to many of the unusual properties of this substance.[1,2] However, a remarkably small fraction of the immense literature on water deals with questions of connectivity—perhaps because there is no experimental probe that directly measures the hydrogen bond distribution.[3] Computer simulation methods, such as molecular dynamics (MD), offer the opportunity to study connectivity in great detail. For example, Geiger *et al.*[4] analyzed the classic MD simulations of Rahman and Stillinger[5,6] from the point of view of connectivity, and found that water is well above its percolation threshold for any reasonable definition of a hydrogen bond.

In this work, we shall extend the work of Ref. 4 by focusing not on questions relating to the percolation threshold, but rather on a family of distribution functions W_M that give the weight fraction of water molecules belonging to finite "nets" of M molecules. These distributions are sufficient to calculate many desired connectivity properties, and thus contain the most general information concerning the hydrogen bond network in liquid water. We shall compare the MD results with calculations based on the concepts and techniques of random-bond percolation theory.[7]

We also consider the connectivity properties of those molecules with four intact bonds. These properties are interesting because a recently proposed structural model regards water to be a transient hydrogen-bonded "gel," local patches of which exhibit a high degree of bondedness.[8] Evidence[9,10] suggests that the local density of such patches is less than the global density, thereby explaining some of the unusual behavior of water. We calculate the distribution functions giving the weight fraction of water molecules belonging to s-molecule patches of four-bonded molecules, and compare the MD findings with calculations based on correlated-site percolation theory.

This work is organized as follows. In Sec. II, we describe the MD simulation approach, and the relation between the definition of a hydrogen bond and the mean number n_{HB} of hydrogen bonds per molecule. In Sec. III we present MD results for f_j, the fraction of water molecules with j intact bonds, and we show the degree to which the MD results are described by a binomial distribution. In Sec. IV, we present MD data for the microscopic distribution functions for M-molecule nets, and compare with the predictions of percolation theory. In Sec. V we perform the analogous calculations for the distribution functions for the fraction of molecules belonging to a s-molecule patch. In Sec. VI we use the microscopic distribution functions to calculate a macroscopic function, the "weight-average" size of the networks. Section VII summarizes the main conclusions of this study. Appendix A presents evidence that regions of *high* bond connectivity are associated with regions of low density. Finally, Appendix B describes the details of the percolation theory calculations.

II. MD SIMULATION AND HYDROGEN BOND DEFINITIONS

The MD system consists of 216 water molecules, interacting via a ST2 pairwise potential, confined to a cubic box with periodic boundary conditions and edge length of 18.6 Å. The system density is 1.0 g/cm^3 and the temperature 284 K. Several hundred equally spaced configurations with $\Delta t = 8.5 \times 10^{-15}$ s of the Stillinger–Rahman simulation run[6] were analyzed.

Because the nature and strength of the interaction among the molecules vary continuously with mutual separation and orientation, any definition that distinguishes unambiguously between broken and intact hydrogen bonds implies drastic simplifications. One possible approach, the energy criterion, was introduced by Rahman and Stillinger.[5,6] Two molecules i and j are considered to be bonded if their interaction energy V_{ij} is stronger (more negative) than some chosen cutoff energy V_{HB}.

For only weakly negative V_{HB}, this purely energetic definition of a bond will also include interactions that are not compatible with accepted conceptions of hydrogen bonds (for example, attractive electrostatic interactions between next-nearest neighbors will also be included). To remedy this defect, we introduced additional restrictions that lead finally to two different definitions of a hydrogen bond:

D_1 (energetic definition/number limit): If more than four bonds per molecule meet the energy criterion, only the four strongest bonds are regarded as intact.

D_2 (hybrid energetic/geometric definition): If any bond meeting the energetic criterion has oxygen–oxygen separation larger than 3.5 Å, it is regarded as broken. This cutoff is roughly the distance of the first minimum in the oxygen–oxygen pair correlation function and is close to the value of 3.4 Å, which has been claimed from x-ray and neutron diffraction studies as a limiting distance for hydrogen bonds.[11] For weakly negative values of V_{HB} we find some molecules with more than four intact bonds.

Ideally, we would like to choose one reasonable value for V_{HB} and then change the temperature of the MD simulation to vary n_{HB}. Unfortunately, it is very time consuming to carry out several simulation runs at different temperatures. By holding the temperature constant and varying V_{HB}, we can approximate the behavior of n_{HB} for conditions of constant V_{HB} and varying temperature. The most negative choice of V_{HB} corresponds to high temperature because it allows few intact bonds; the least negative choice of V_{HB} corresponds to low temperature because it allows many intact bonds. Since the choice of V_{HB} is arbitrary, we can vary V_{HB} gradually to obtain a systematic "scan" of the continuous range of bond energies.

We vary the parameter V_{HB} over a sequence of 32 discrete values $V_{HB} = -82\epsilon, -80\epsilon, -78\epsilon,..., -20\epsilon$, where $\epsilon = 0.075\ 75$ kcal/mol. As V_{HB} becomes less negative we say it becomes more permissive, because the mean number of hydrogen bonds per molecules n_{HB} increases. This variation is shown in Fig. 1 for both bond definitions. The curves are roughly the same for both D_1 and D_2, except for very permissive choices of V_{HB}, where D_2 leads to larger values of n_{HB} because it allows more than four bonds.

Our results are rather general in the sense that they are not very sensitive to the details of the choice of the hydrogen bond definition. Indeed, it appears that connectivity properties are comparable (i) for energetic and geometric definitions, (ii) for different simulation methods (MD or Monte Carlo), and (iii) for different molecular interaction models (e.g., ST2 or MCY).[12-16] For example, some of the results of this work for the distribution functions $W_1 - W_4$ have been confirmed using a completely different simulation method: a Monte Carlo study with a MCY potential using a geometric definition of a hydrogen bond.[17]

III. FRACTION OF MOLECULES WITH j INTACT BONDS

First we direct our attention to the water molecules themselves, and in Secs. IV–VI below we consider their network properties. We begin by calculating f_j, the fraction of water molecules with j intact bonds, which is related to the total number of bonds per molecule by

$$n_{HB} = \sum_{j \geq 1} j f_j. \tag{1}$$

For purposes of making graphs, we shall henceforth use n_{HB} as the independent variable, or the number $p = n_{HB}/z$, which can be interpreted roughly as the fraction of intact bonds (here z is the functionality of a water monomer). In this fashion, we omit the auxiliary variable V_{HB}, since n_{HB} is the more essential parameter for the description of bond networks, allowing a coincident description of systems in very different states.[4] Figure 2(a) shows the f_j as a function of $n_{HB} = zp$ for both definitions of a hydrogen bond. The solid curves are the simple binomial distribution
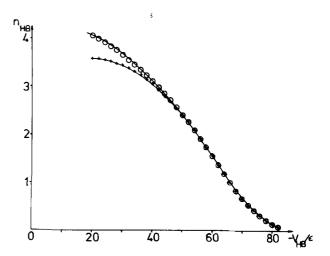
$$f_j = \binom{z}{j} p^j (1-p)^{z-j}, \tag{2}$$

FIG. 1. Dependence of n_{HB}, the average number of hydrogen bonds per water molecule, on the cutoff energy V_{HB} for bond definitions $D_1 (+)$ and $D_2 (\bigcirc)$ ($\epsilon = 0.075\ 75$ kcal/mol). The two definitions produce remarkably similar results, different only in the range of very permissive V_{HB}, where we find many five-bonded molecules using definition D_2 and none using D_1 (which allows a maximum of four bonds per molecule).

where $z = 4$ in Fig. 2(a) and $z = 5$ in Fig. 2(b). Remarkably the MD points agree with the $z = 4$ curve over a very large range of p. There are strong deviations only for values of p larger than about 0.8, corresponding to very permissive definitions of a hydrogen bond. In this region, we find a substantial fraction of molecules with five intact bonds when using definition D_2. Nonetheless, the data agree better with the $z = 4$ curves than with the $z = 5$ curves.

The agreement between the MD data and the binomial formula shows that the approximation of statistically independent bond formation is a sound first order approach, at least in describing "ST2 water" which assumes a pairwise additive potential function, excluding many-body forces. We shall see in Sec. V that there is some evidence for a weak cooperative effect in the distribution functions for the patches of four-bonded molecules.

IV. THE BOND NETWORK DISTRIBUTION FUNCTIONS $W_M(p)$

The hydrogen bond network analysis of the MD configurations was made along the lines described in Ref. 4 (where a purely energetic bond definition was used exclusively). For
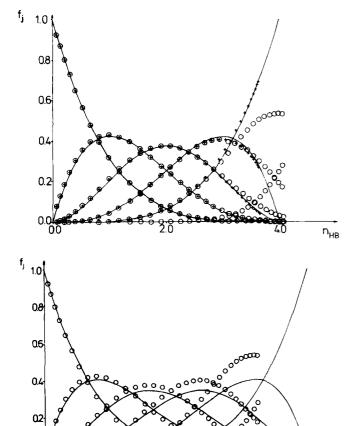

FIG. 2. (a) Fraction f_j of water molecules having exactly j intact hydrogen bonds ($j = 0,1,...,5$) calculated using bond definitions D_1 (+) and D_2(O). The full lines correspond to the binomial distribution [Eq. (2)] with $p = n_{HB}/4$ and $z = 4$. (b) The f_j for definition D_2 compared with the binomial distribution for $z = 5$.

a fixed value of V_{HB} which corresponds to a fixed value of $p = n_{HB}/4$, we calculate the average number $n_M(p)$ of nets of M molecules. Hence the weight fraction of molecules belonging to such "M nets" is simply

$$W_M(p) = M n_M(p)/216, \qquad (3)$$

since there are 216 molecules in the system. Note that $W_M(p)$ is the probability that a randomly chosen molecule will be connected by a path of hydrogen bonds to $M - 1$ other molecules.

The actual calculation of $n_M(p)$ was described in Appendix A of Ref. 4. Now, however, we omit all nets that span the basic 18.6 Å cube in at least one coordinate direction as these are indistinguishable from infinite nets. In our system, this procedure has no effect for $M < 20$, but we shall see that it leads to a sharply peaked maximum in functions such as the average molecular weight (cf. Sec. VI below).

In any analysis using percolation theory, it is essential to first eliminate all connected nets that span the finite dimension of the sample, since percolation theory concerns only finite clusters. To identify spanning nets, the following procedure has been used. First, we identify all the pairs of water molecules that belong to the net in consideration and that are connected in the periodic arrangement across a face of the basic cube. Second, starting from one pair member, we search for the other pair member by following all paths along bonds that do not cross the separating face. If we can connect any such pair in the net by this method, it is said to span.

In total, 127 MD configurations were analyzed for bond definition D_1 and 800 configurations for bond definition D_2. Hence we expect the results for D_2 to show less statistical fluctuation. We obtained 50 graphs of $W_M(p)$ against p using both D_1 and D_2. Figure 3 shows examples of these plots for the cases $M = 2$, 3, 5, and 12; note that the special case $M = 1$ corresponds to isolated monomers so that $W_1(p)$ is the same function as f_0 plotted in Fig. 2. Also shown in Fig. 3 are the results of analytic random bond percolation calculations. These were carried out for the ice lattice, since this lattice perhaps best captures the essential features of the local environment of a water molecule. We emphasize, however, that the bond percolation calculations are very insensitive to the lattice chosen—indeed, calculations for a Cayley tree give essentially the same curves for $M < 30$ (Ref. 18) (a Cayley tree pseudolattice assumes that each water monomer can interact with up to z other monomers, that the monomers are not constrained to a lattice, and that cycles have a negligible effect on the distribution functions). The percolation calculations for the ice lattice are carried out analytically for $M \leqslant 6$ (see Appendix B), and the exact expressions are given in Table I. For $M > 6$, the analytic calculations become prohibitively complex due to the immense number of possible configurations.

We see in Fig. 3 excellent agreement between the MD results from the two different bond definitions. Even more remarkable is the astonishing agreement between the MD simulation results and the percolation theory calculations which are carried out on the assumption of random bond formation. This agreement—with no adjustable parameters—may be explained by the fact that the observed nets
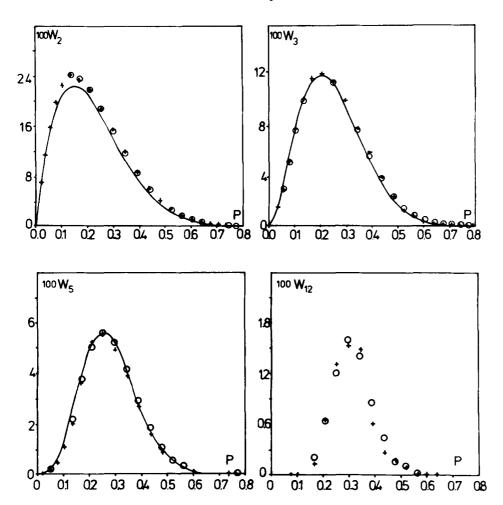
FIG. 3. MD results for $W_M(p)$, the weight fraction of molecules belonging to nets of size $M = 2$, 3, 5, and 12 for definitions $D_1(+)$ and $D_2(O)$. The solid curves for $M = 2, 3$, and 5 are the random-bond percolation calculations described in Appendix B. No adjustable parameters are used, so the agreement is rather remarkable.

are still relatively small, so that the local tetrahedral order is the most decisive factor. This explanation is supported by recent work showing that fair agreement is obtained when the lattice calculations are replaced with Cayley tree calculations.[18]

V. "PATCHES" OF FOUR-BONDED WATER MOLECULES

In the previous section we considered the "transient gel" or infinite network formed by hydrogen bonds linking together individual water molecules. By itself, this network is not sufficient to explain the unusual behavior of H_2O and D_2O. For example, we need some mechanism that leads to increased density fluctuations if we are to explain the unusually large isothermal compressibility observed. Recently it was proposed[8] that the local density and entropy varied from point to point in the hydrogen bond network, and that these anomalous variations were correlated with the degree of bondedness of the water molecules. It is convenient but not necessary to focus *mainly* on those water molecules with four intact bonds; these form tiny regions or "patches" within the gel network. It was suggested[8] that the local density and local entropy of these patches is less than the global density and global entropy of the entire gel, thereby giving

TABLE I. Weight fraction W_M of molecules belonging to M-molecule bond networks, as a function of the random bond probability p. The calculations are carried out for an ice I_h lattice and are exact to order $M = 6$. The case $M = 1$ is identical with f_0 of Eq. (2). A useful check on the accuracy of the calculations is the fact that the sum on M of W_M must equal the total weight fraction of all molecules belonging to networks of any size—so that the coefficients of all successive powers of p must vanish identically. Also given, for comparison, is the exact result W_M^{CT} for a "Cayley tree," which corresponds to neglecting all networks containing closed loops; this is the Flory theory (see Ref. 18).

$$W_1 = (1 - p)^4$$

$$W_2 = 4p(1 - p)^6$$

$$W_3 = 18p^2(1 - p)^8$$

$$W_4 = 88p^3(1 - p)^{10}$$

$$W_5 = 455p^4(1 - p)^{12}$$

$$W_6 = 12p^5(1 - p)^{12}[198(1 - p)^2 + 5(1 - p) + 1]$$

...

...

...

$$W_M^{CT} = MA(M)p^{M-1}(1 - p)^{2M + 2} \text{ (valid for all } M \text{ on Cayley tree, "Flory theory"),}$$
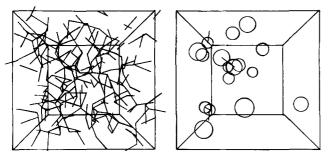
where

$$A(M) = 4(3M)!/M!(2M + 2)!$$

FIG. 4. Results of an instantaneous snapshot of the MD simulation of 216 water molecules constrained to a cube of edge 18.6 Å (shown in perspective). Part (a) shows only intact bonds, as defined by definition D_1 with cutoff parameter $V_{HB} = -56\epsilon = -4.2$ kcal/mol, corresponding to $n_{HB} = 1.92$ or $p = 0.48$, well above the bond percolation threshold. In part (b) are shown as large open circles those molecules with four intact bonds.

rise to a physical mechanism for the observed anomalies in the isothermal compressibility (density fluctuations), constant-pressure specific heat (entropy fluctuations), and thermal expansivity (coupled density/entropy fluctuations). These suggestions have recently received support from the small-angle x-ray scattering data of Bosio *et al.*[10] which display Lorentzian behavior with characteristic length scale of 8 Å, coincident with the characteristic length scale of the four-bonded patches calculated by MD. Moreover, in these calculations the local density in ST2 water has been probed and found to be decreased in the vicinity of the patches.[9]

For these reasons, a detailed connectivity analysis of the four-bonded molecules was carried out. For typical realistic values of V_{HB}, the bond nets are well above the bond percolation threshold (so that virtually all of the water molecules belong to the infinite cluster), but the patches of four-bonded molecules can still be relatively small. This fact is illustrated in Fig. 4. Part (a) is an instantaneous snapshot of the system showing the bond networks, while part (b) shows by a circle each of the four-bonded molecules.

To study the distribution functions $W_s^*(p)$ characterizing patches or clusters of four-bonded molecules, we analyzed the MD simulations in an analogous fashion to that used above for bond nets. When in definition D_2 there occurs a molecule with five intact bonds, it is also considered to belong to the patch.

Figure 5 shows the distribution function $W_1^*(p)$ for a "one-molecule patch"—i.e., a single four-bonded molecule whose neighbors are not four bonded, while Fig. 6 shows $W_s^*(p)$ for $s = 2, 3, 5$, and 12 in analogy to Fig. 3. Again, excellent internal consistency between bond definitions D_1 and D_2 is found. The statistical accuracy decreases with increasing s, because the absolute number of patches decreases. Near the maxima, the statistical error increases from 5% for $s = 1$ to about 30% for $s = 12$ for definition D_1, while for D_2 the error increases from about 2% for $s = 1$ to about 10% for $s = 12$.

The continuous curves in Figs. 5 and 6 give the predictions of correlated-site percolation theory $W_s^c(p)$. For $s \leqslant 6$,
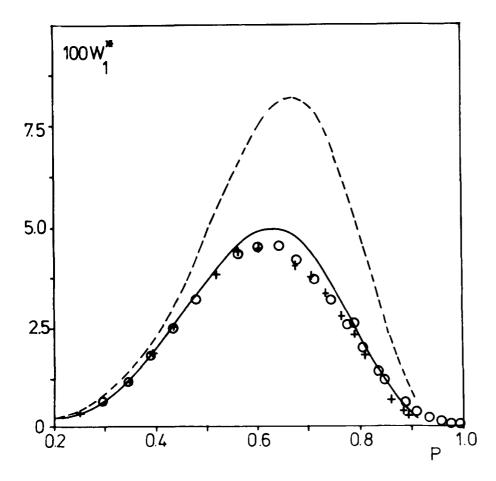

FIG. 5. MD results for $W_s^*(p)$, the weight fraction of molecules belonging to patches (clusters of four-bonded molecules) of size $s = 1$ for definitions $D_1(+)$ and $D_2(O)$. Also shown are the predictions of correlated site percolation (full line) and random-site percolation (dashed line).

the calculations have been carried out analytically using the methods described in Appendix B; these exact results are given in Table II.

For larger values of s, Monte Carlo simulations were carried out for a finite lattice of 21 000 molecules; these reproduce the analytic results for $s \leqslant 6$ and are believed reliable for $s > 6$. The statistical accuracy of these calculations was found to be better than 1%. To test whether the correlation was responsible for the agreement, we also considered the cluster distribution functions W_s^R for randomly distributed sites. Analytical results of W_s^R for the ice lattice are not available, but for the closely related diamond lattice they are known up to $s = 12$.[19] These continuous curves are shown dashed in Figs. 5 and 6. the MD results clearly agree much better with W_s^C. Thus it is not percolation concepts in general but correlated site percolation in particular that is relevant to the MD data.

We note that for $s = 1$, $W_1^C(p) < W_1^R(p)$ over the entire range of p. This result reflects the fact that in correlated site percolation, sites tend to "clump" together more than in random site percolation. Thus the probability of finding small clusters diminishes. With increasing values of s, we also observe the opposite behavior: $W_s^C > W_s^R$—at first only for small values of p and then for larger values, thus leading to an intersection of the curves $W_s^C(p)$ and $W_s^R(p)$. By this

the initial "excess" for small s is compensated at larger values of s, because the sum over all probabilities at a fixed value of p must be constant and equal to one.

As stated above, the MD data agree much better with the results of correlated site percolation theory than with random site percolation theory. On a finer scale, we note an interesting systematic behavior: the small deviations of W_s^* from W_s^C are always in the opposite direction compared to the random site probabilities W_s^R. In other words, when $W_s^C < W_s^R$, the MD data W_s^* are even smaller than W_s^C, and in the opposite case even larger. The intersection between the W_s^C and W_s^* curves is also near the previously discussed intersection between W_s^C and W_s^R.

This behavior can be interpreted as the manifestation of a weak additional *cooperative* effect, which increases the tendency of clumping together beyond the pure statistical correlation of sites. Of course, the underlying mechanism cannot be reinforcement of hydrogen bonds due to many-body interactions because the simulation is done on the basis of pairwise additive forces, but is probably a weak mutual stabilization of four-bonded water molecules, comparable to the mutual stabilization of larger structural entities, like water octomers, as discussed by Stillinger.[20,21] Also the observation of increased hydrogen bonding in the vicinity of hydrophobic particles seems to be a related effect.[22]
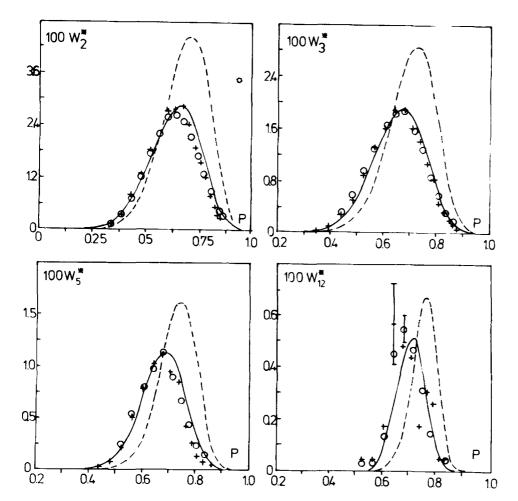

FIG. 6. Same as Fig. 5, except that $s = 2$, 3, 5, and 12. The agreement between the MD results and correlated-site percolation theory is quite striking.

TABLE II. Exact ice lattice expressions for W_s^* for the weight fraction of water molecules belonging to a patch of s molecules. Also given is the Flory theory expression (Ref. 18), where $A(s)$ is defined in Table I.

$$W_1^* = p^4(1 - p^3)^4$$

$$W_2^* = 4p^7(1 - p^3)^6$$

$$W_3^* = 18p^{10}(1 - p^3)^8$$

$$W_4^* = p^{13}[46(1 - p^3)^{10} + 36(1 - p^3)^8(1 - 2p^3 + p^5) + 6(1 - p^3)^6(1 - 2p^3 + p^5)]$$

$$W_5^* = p^{16}[63(1 - p^3)^{12} + 208(1 - p^3)^{10}(1 - 2p^3 + p^5)$$

$$+ 107(1 - p^3)^8(1 - 2p^3 + p^5)^2 + 2(1 - p^3)^6(1 - 2p^3 + p^5)^3$$

$$+ 15(1 - p^3)^9(1 - 3p^3 + 2p^5 + p^6 - p^7) + 30(1 - p^3)^{10}(1 - p^2)$$

$$+ 30(1 - p^3)^8(1 - p^2)(1 - 2p^3 + p^5)]$$

$$W_6^* = 9p^{18}(1 - p^3)^{10}(1 - 2p^3 + p^5)$$

$$+ 3p^{18}(1 - p^3)^{12}$$

$$+ 24p^{19}(1 - p^3)^6(1 - 2p^3 + p^5)^4$$

$$+ 4p^{19}(1 - p^3)^6(1 - p^2)(1 - 2p^3 + p^5)^3$$

$$+ 40p^{19}(1 - p^3)^9(1 - p^2)(1 - 3p^3 + 2p^5 + p^6 - p^7)$$

$$+ 18p^{19}(1 - p^3)^8(1 - 3p^3 + 2p^5 + p^6 - p^7)^2$$

$$+ 239p^{19}(1 - p^3)^8(1 - 2p^3 + p^5)^3$$

$$+ 134p^{19}(1 - p^3)^8(1 - p^2)(1 - 2p^3 + p^5)^2$$

$$+ 142p^{19}(1 - p^3)^9(1 - 2p^3 + p^5)(1 - 3p^3 + 2p^5 + p^6 - p^7)$$

$$+ 637p^{19}(1 - p^3)^{10}(1 - 2p^3 + p^5)^2$$

$$+ 331p^{19}(1 - p^3)^{10}(1 - p^2)(1 - 2p^3 + p^5)$$

$$+ 36p^{19}(1 - p^3)^{10}(1 - p^2)^2$$

$$+ 18p^{19}(1 - p^3)^{11}(4p - 7p^2 + 3p^3)$$

$$+ 61p^{19}(1 - p^3)^{11}(1 - 2p^3 + p^5)(1 - 3p^3 + 2p^5 + p^6 - p^7)$$

$$+ 440p^{19}(1 - p^3)^{12} + 173p^{19}(1 - p^3)^{12}(1 - p^2) + 79p^{19}(1 - p^3)^{14}$$

$$W_s^* \text{ [Flory theory]} = sA(s)p^{3s + 1}(1 - p^3)^{2s + 2}$$

VI. AVERAGE NETWORK SIZES

From the complete distribution of bond net sizes $W_M(p)$ for all M, one can calculate the average size of the finite bond nets,[23]

$$S(p) = \sum_{M > 1} M W_M(p). \tag{4}$$

The corresponding average cluster size $S^*(p)$ can be obtained from the distribution functions $W_s^*(p)$. Both $S(p)$ and $S^*(p)$ are shown in Fig. 7 for both bond definitions.

Percolation theory predicts for our finite system a sharp peaking of these functions at the respective percolation thresholds p_c and p_c^*. The values of n_{HB} where the maxima occur are in perfect agreement with the corresponding critical values obtained from the Monte Carlo simulations on the ice lattice, $p_c = 0.388$ and $p_c^* = 0.795$. Not only are the threshold values for the MD data the same as in percolation theory, but even the critical exponents describing the singularity in S (and in S^*) are the same.[24]

These two percolation points have played some role in recent discussions on the properties of liquid water. As had been conjectured in one of the early "gel model" approaches[25] and confirmed by MD calculations,[4] liquid water is always above the bond percolation threshold—at $n_{HB} \cong 1.53$. On the other hand, in an attempt to explain the unusual properties of water, it was proposed[8] that water was below the percolation threshold of the four-bonded molecules—at $n_{HB} \cong 3.18$. Combining these two pictures, one may state that the hydrogen bond connectivity of liquid water is somewhere in the region limited by the positions of the two maxima in Fig. 7.

VII. DISCUSSION AND SUMMARY

In this work we have tried to answer the question of how to describe hydrogen-bonded aggregates of water molecules; more generally, we have examined how a local interaction, the hydrogen bond, leads to global properties of the macroscopic network we call water. This question is addressed, at least indirectly, by all of the numerous structural models of water that have been proposed during recent decades.

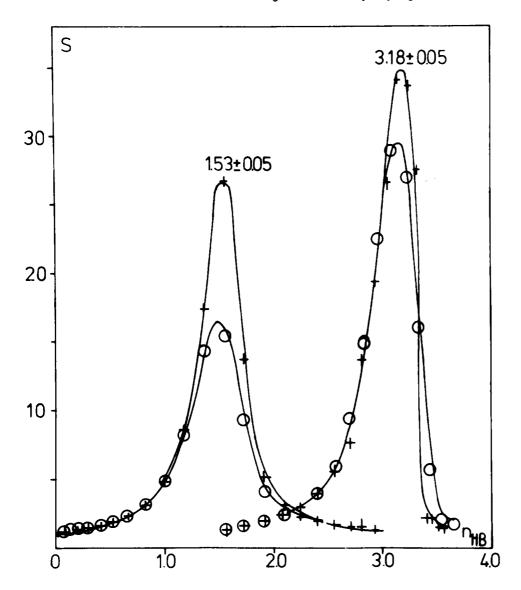The following picture emerges from our analysis of the

FIG. 7. Peaks at $n_{HB} = 1.53$: MD results for the average sizes of finite bond networks for bond definitions $D_1(+)$ and $D_2(O)$; peaks at $n_{HB} = 3.18$: average size s^* of the finite patches of four-bonded molecules for $D_1(+)$ and $D_2(O)$. The value of n_{HB} corresponding to "real water" is presumably somewhere in between the two peaks, so that water is above the bond percolation threshold but below the percolation threshold of the patches of four-bonded molecules.

ST2 MD calculations: water appears to be a macroscopic, space-filling gel-like network[4,8,25] held together by relatively short-range microscopic hydrogen bonds between neighboring molecules; the network contains small regions (patches) of increased connectivity, and associated with these patches is a decreased local density. This structural picture of water combines ideas of continuum models (the extensive distorted hydrogen-bond network) with older ideas of mixture models (small regions with different local order and density). These two views of water are traditionally considered mutually exclusive. We have seen that depending on which questions we ask, either viewpoint can emerge: Questions of hydrogen bond networks require a continuum picture while questions concerning locally structured regions require the structured patches provided by the statistics of the four-bonded molecules. For example, measurements like neutron scattering experiments, designed to investigate the average microscopic structure of liquid water, will favor the continuum picture, whereas the measurement of fluctuation-determined thermodynamic properties like the compressibility or heat capacity will be explained by the presence of the structured patches.

Perhaps the most striking feature of the above analysis is the degree of quantitative agreement—without adjustable parameters—between the MD simulations on the one hand and the predictions of percolation theory on the other. Representative examples of this agreement are shown in Figs. 2, 3, 5, 6, and 7. The initial assumption of random bond formation underlies all the percolation calculations. Thus the agreement with percolation theory of the microscopic distributions functions W_M and W_s^*, as well as the macroscopic parameters like the mean cluster size, the percolation threshold, and even the exponents characterizing the singular behavior at the percolation threshold strongly suggests that on the time scale of bond formation these bonds form randomly.

But even in our simple case of strict pair additive interactions, we cannot claim that there is absolutely no "cooperativity" in bond formation. Indeed, small but apparently systematic deviations from the cluster distributions obtained from percolation theory applying the random bond assumptions *may* reflect the presence of a weak cooperative effect, which leads—by mutual stabilization of four-bonded molecules—to an increased formation of larger clusters of four-

bonded molecules. Additional MD calculations are underway to study the possible increase of this effect at lower temperatures. In "real" water, a contribution to the tendency of clumping together might be expected due to the presence of many-body interactions[26] which are not explicitly included in the ST2 simulations.

Furthermore, the present studies refer exclusively to "snapshots" of the hydrogen-bond network and do not contain any information about the dynamics of the system. Since the "transient gel" of liquid water is distinguished from a more familiar "chemical gel" by the picosecond lifetime of the bonds, the network is continually restructuring itself. Obviously there must be relations between the mean lifetime of single bonds, the lifetime of larger aggregates, and the magnitude of transport coefficients. There have not been many attempts to elucidate these connections.[27] Since in the supercooled region singular behavior of transport coefficients has been observed, the results of the present studies should be supplemented by a study of the dynamic properties of the hydrogen bond network, and these studies are underway.

ACKNOWLEDGMENTS

APPENDIX A: CORRELATION BETWEEN CONNECTIVITY AND "LOCAL" DENSITY

To show that decreased local density occurs within patches of increased connectivity, we discuss the results shown in Fig. 8. In these graphs we correlate the number n of neighbors j, which we find within a sphere of radius r_c around some reference molecule i, with the sum u_i of the corresponding pair interaction energies v_{ij},

$$u_i = \sum_{j \neq i}^{n} v_{ij}. \tag{A1}$$

Here u_i is the binding energy of the reference molecule i with respect to its n neighbors in the sphere of radius r_c; it is a measure of the local connectivity, which is more general than the hydrogen bond picture and which avoids the arbitrary definitions of a hydrogen bond D_1 and D_2.

Due to the fluctuation of the local density, we observe a range of numbers n. We now calculate the average binding energy as a function of the number of neighbors n

$$u = \langle u_i \rangle_{n = \mathrm{const}}. \tag{A2}$$

These values are shown in Fig. 8 for four different choices r_c; the vertical bars indicate the mean square deviations from the averages u.

These deviations are smallest in the central part of the graphs, because those numbers of neighbors n we find most frequently and therefore we have many contributions to the
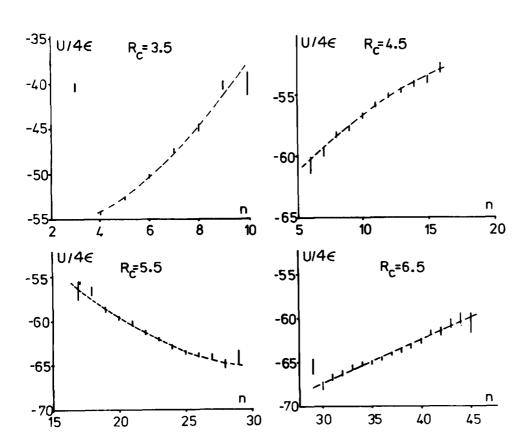

FIG. 8. Correlation between the number n of neighbors found within a sphere of radius r_c around some reference molecule and the corresponding building energy u of this molecule ($\epsilon = 0.075\,75$ kcal/mol). We find "normal" *decrease* of u with n, when choosing $r_c = 5.5$ Å, but "typically water-like" *increase* for $r_c = 3.5$, 4.5, and 6.5 Å.

corresponding averages. Through these more reliable points a dashed guidance line has been drawn. Averages from less than 100 contributions, which occur at the outermost wings of the distributions (very high and very low n) are not considered in these graphs.

We see that for $r_c = 5.5$ Å the average binding energy u decreases with an increasing number of neighbors within the sphere of regarded interaction. This is what we would expect from a "normal" liquid like a Lennard-Jones liquid at not too high packing densities, because the addition of another interaction partner will add a negative (attractive) contribution v_{ij}. However, in the cases $r_c = 3.5$, 4.5, and 6.5 Å, we observe exactly the opposite behavior; u increases with increasing local density. This means that a less dense local arrangement of the water molecules is energetically favorable over more dense structures; a behavior that we regard as typically "water-like" is related to the occurrence of the anomalies.

The observation that for some choices of r_c we get the picture of a normal liquid has already been reported in a previous paper and can be explained by the oscillatory nature of the pair correlation functions, which describe the local structure of water. Figure 2 of Ref. 9 indicates a decreased local density around four-bonded water molecules when using $r_c = 3.5$, 4.5 or 6.5 Å, whereas for the choice $r_c = 5.5$ Å no such difference can be observed.

Thus the present results shown in Fig. 8 confirm and also generalize our previous finding of a correlation between increased connectivity and decreased local density.

Furthermore, concentrating on the graph for $r_c = 3.5$ Å (a value which had been used before as the limiting distance for hydrogen bonds), Fig. 8 indicates a marked minimum of u at $n = 4$. This indicates again a strong energetic preference for four-coordinated local structures.

APPENDIX B: PERCOLATION THEORY APPLIED TO HYDROGEN BOND NETWORKS

Here we adapt conventional random bond percolation to calculate the microscopic distribution functions $W_M(p)$ (for nets of water molecules) and $W_s^*(p)$ (for patches of four-bonded water molecules).
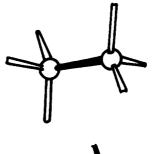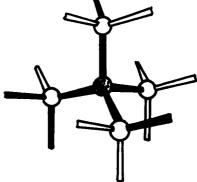
We begin with two important assumptions:

(i) Bond formation in liquid water is more or less random on the time scale of picoseconds.

(ii) A lattice may reflect to some degree the local structure of hydrogen bonds in water. The similar ice and diamond lattices as well as the Cayley tree pseudolattice are all possible choices.

For any lattice we choose, assumption (ii) breaks down at some characteristic length scale. Of ice, diamond, and Cayley tree, the ice lattice agreed with the MD results to the largest length scale, about $M = 50$. We emphasize, however, that there must certainly be some limit to this surprising agreement; water does not have the same connectivity as ice. We also recognize that the ST2 potential used in our simulation has been criticized for being too directional, a fault which might lead to an excessively ice-like hydrogen-bond structure. In any case, the ice lattice seems to be a good choice, though diamond might be just as good.

Consider an infinitely large ice lattice with a fraction p of its bonds intact and a fraction $(1 - p)$ of its bonds broken at random. The functions $W_M(p)$ and $W_s^*(p)$ which we have described as weight fractions can now be thought of as probabilities. For instance, the weight fraction of nets of water molecules of size $M = 2$ is simply the probability that a randomly chosen water molecule belongs to such a net. To do so it must satisfy several conditions: it must have one bond connecting it to another molecule; all its three other bonds must be broken; and all of its partner's three other bonds must be broken. Such a configuration is shown in Fig. 9(a). Thus we require a total of one intact bond and six necessarily broken bonds, and the probability of this configuration is just $p(1 - p)^6$. On the ice lattice there are four possible orientations for size-2 nets, hence

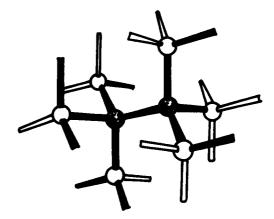$$W_{M=2}(p) = 4p(1 - p)^6. \tag{B1}$$

FIG. 9. Configurational examples to illustrate the calculation of cluster probabilities (a) $W_{M=2}(p)$, (b) $W_{s=1}^*(p)$, (c) $W_{s=2}^*(p)$. Intact and broken bonds are indicated by full and open bars.

To continue this process for larger nets, we enumerate all possible configurations of the M molecules, keeping track of the number of necessarily intact and broken bonds in each as well as the number of possible orientations. The resulting $W_M(p)$ is

$$W_M(p) = \sum_{n=1}^{N} \text{(number of orientations)}\, p^{i(n)}(1-p)^{b(n)},$$
(B2)

where there are N configurations, and $i(n)$ intact and $b(n)$ broken bonds in configuration n. The limiting factor in this calculation is the calculation time to enumerate the large number of configurations possible for large M. The exact formulas for $M = 1$ to $M = 6$ are given in Table I.

The calculation of $W_s^*(p)$ is slightly more complicated because we must consider a large number of bonds even for small patches. For instance, consider $W_s^*(p)$ for $s = 1$, the weight fraction of isolated four-bonded molecules. This is the probability that a randomly chosen molecule will be four bonded and that none of its neighbors will be four bonded, a condition which involves 16 bonds. We show such a configuration in Fig. 9(b). It is immediately clear that this bond configuration is by no means the only one that leads to an isolated four-bonded molecule. The neighboring molecules may have either 1, 2 or 3 intact bonds each—it makes no difference, as long as they do not have 4. The easiest way to find the probability that a neighbor site is *not* four bonded is to find the probability that it *is* four bonded and subtract. If the center molecule is already four bonded, then each neighbor molecule only needs three more bonds to be four bonded. Thus the probability that each neighbor molecule is *not* four bonded is $(1 - p^3)$. In this way we find

$$W_{s=1}^*(p) = p^4(1 - p^3)^4.$$
(B3)

We refer to the $(1 - p^3)$ factors as "perimeter" terms, terms that describe the necessary conditions of the molecules immediately next to the molecules in the patch itself. For $W_{s=2}^*(p)$ the perimeter terms are still relatively simple. From Fig. 9(c) we see that there are seven bonds which must be intact and six perimeter molecules which must not be four bonded. Noting that there are four possible orientations of an $s = 2$ patch on an ice lattice, we have

$$W_{s=2}^*(p) = 4p^7(1 - p^3)^6.$$
(B4)

Extending this process to larger patches is difficult because of the increasing complexity and variety of the perimeter terms. Complications arise when a single perimeter molecule shares bonds with more than one molecule in the patch, and when perimeter molecules share bonds with each other. In general we find

$$W_s^*(p) = \sum_{n=1}^{N} \text{(number of orientations)}$$
$$\times p^{i(n)} \text{(perimeter terms)},$$
(B5)

where there are N configurations with $i(n)$ intact bonds in configuration n. In practice, we determine perimeter terms for a certain number of commonly found configurations of perimeter molecules; then we enumerate all possible configurations of the s-molecule patch and piece together the perimeter terms as we recognize familiar patterns in the neigh-

bor molecules. Exact results through $s = 6$ are shown in Table II. This enumeration and recognition process is carried out by computer, but those perimeters which the algorithm does not recognize are left for individual consideration. With this method the limiting factor is the increasing complexity necessary in the computer program, whereas in the $W_M(p)$ calculation the limiting factor was only calculation time. We could overcome the programming problem by enumerating all possible bond configurations for each s-site patch, but there is not much to be gained by this method since the number of bond configurations grows very quickly for even small s ($\simeq 2^{12}$ possible for $s = 1$).

For large M and large s we turn to another method, a Monte Carlo calculation. Imagine a large but finite ice lattice. Using a random number assign intact bonds with probability p. Then we count the nets or patches in the resulting lattice using an algorithm originally described by Hoshen and Kopelman.[28] To determine the weight fractions $W_M(p)$ and $W_s^*(p)$ it is desirable to use as large a lattice as possible, considering memory and computing time limitations, and to perform as many realizations as possible at a variety of values of p. Because we must use discrete values of p, the curves we find are really a series of points. These curves are approximations to the exact functions that we derive by the method of enumeration described previously, and they reproduce the exact solutions very well. In addition, the Monte Carlo method can be used to determine approximations for much larger M and s, the only limitations being lattice size and computation time. A typical large lattice for our Monte Carlo calculations contains $22 \times 22 \times 44$ sites and $2 \times 22 \times 22 \times 44$ bonds, with at least 200 realizations for each value of p. Under these conditions the statistical uncertainty for W_s^* appears to be reasonably small up to at least $s = 50$.

[1](a) L. Pauling, *The Nature of the Chemical Bond* (Cornell University, Ithaca, 1960), Chap. 12; (b) *The Hydrogen Bond*, edited by P. Schuster, G. Zundel, and C. Sandorfy (North–Holland, Amsterdam, 1976).

[2]H. Frank, in *Water*, edited by F. Franks (Plenum, New York, 1972), Vol. 1; W. A. P. Luck, Angew. Chem. Int. Ed. Engl. **19**, 28 (1980).

[3]C. A. Angell, Annu. Rev. Phys. Chem. **34**, 593 (1983) and references therein.

[4]A. Geiger, F. H. Stillinger, and A. Rahman, J. Chem. Phys. **70**, 4185 (1979).

[5]A. Rahman and F. H. Stillinger, J. Chem. Phys. **55**, 3336 (1971).

[6]F. H. Stillinger and A. Rahman, J. Chem. Phys. **60**, 1545 (1974).

[7]D. Stauffer, Phys. Rep. **54**, 1 (1979); for applications to gelation, see D. Stauffer, A. Coniglio, and M. Adam, Adv. Polym. Sci. **44**, 103 (1982).

[8]H. E. Stanley, J. Phys. A **12**, L329 (1979); H. E. Stanley and J. Teixeira, J. Chem. Phys. **73**, 3404 (1980).

[9]A. Geiger and H. E. Stanley, Phys. Rev. Lett. **49**, 1749 (1982).

[10]L. Bosio, J. Teixeira, and H. E. Stanley, Phys. Rev. Lett. **46**, 597 (1981).

[11]I. Olovsson and P. G. Jönsson, in Ref. 1(b), Vol. 2.

[12]M. Mezei and D. L. Beveridge, J. Chem. Phys. **74**, 622 (1981).

[13]W. L. Jorgensen, Chem. Phys. Lett. **70**, 326 (1980).

[14]C. Pangali, M. Rao, and B. J. Berne, Mol. Phys. **40**, 661 (1980).

[15]K. Okazaki, S. Nose, Y. Kataoka, and T. Yamamoto, J. Chem. Phys. **75**, 5864 (1981).

[16]Y. Kataoka, H. Hamada, S. Nose, and T. Yamamoto, J. Chem. Phys. **77**, 5699 (1982).

[17]M. Mezei (private communication).

[18]H. E. Stanley, R. L. Blumberg, and A. Geiger, Phys. Rev. B **28**, 1626 (1983).

[19]M. F. Sykes, D. S. Gaunt, and M. Glen, J. Phys. A **9**, 1705 (1976).

[20]F. H. Stillinger, Science **209**, 541 (1980).

[21]F. H. Stillinger, in *Water in Polymers*, edited by S. P. Rowland (American Chemical Society, Washington, D.C., 1979).

[22]A. Geiger, A. Rahman, and F. H. Stillinger, J. Chem. Phys. **70**, 263 (1979).

[23]The corresponding averages, *including the spanning nets*, were used in Ref. 4.

[24]A. Geiger and H. E. Stanley, Phys. Rev. Lett. **49**, 1895 (1982).

[25]J. H. Gibbs, C. Cohen, P. D. Fleming III, and H. Porosoff, J. Sol. Chem. **2**, 179 (1973).

[26]D. Hankins, J. W. Moskovitz, and F. H. Stillinger, J. Chem. Phys. **59**, 995 (1973); **53**, 4544 (1970).

[27]One such attempt to explain self-diffusion coefficients from the presence of interaction aggregates may be found in H. G. Hertz, Ber. Bunsenges. Phys. Chem. **74**, 666 (1970); **75**, 183 (1971); **75**, 572 (1971).

[28]J. Hoshen and R. Kopelman, Phys. Rev. B **14**, 3438 (1976).