Criticality and Parallelism in Combinatorial Optimization

William G. Macready, Athanassios G. Siapas,* Stuart A. Kauffman

Local search methods constitute one of the most successful approaches to solving large-scale combinatorial optimization problems. As these methods are increasingly parallelized, optimization performance initially improves but then abruptly degrades to no better than that of random search beyond a certain point. The existence of this transition is demonstrated for a family of generalized spin-glass models and the traveling salesman problem. Finite-size scaling is used to characterize size-dependent effects near the transition, and analytical insight is obtained through a mean-field approximation.

Optimization tasks are common and very often uncommonly difficult. In most areas of science and engineering, from free-energy minimization in physics to profit maximization in economics, the need to optimize is ubiquitous. Because of the importance and difficulty of optimization, much effort has gone into developing effective algorithms for finding good optima. The methods of simulated annealing (*1*), genetic algorithms (*2*), and taboo search (*3*) are three of the most popular techniques, inspired by ideas from statistical mechanics, evolutionary biology, and operations research, respectively. All of these methods rely in part on constructing improved solutions by applying a local operator to a population of candidate solutions. Good solutions result from the accumulation of many beneficial local modifications applied one after another. An obvious speed up in the algorithm's performance can be gained if we apply many local modifications in parallel. Despite the promise that parallel algorithms hold, they have received less attention and little is known about them (*4*).

In this report, we investigate the effects of parallelizing local search for combinatorial optimization. We demonstrate that for a wide class of search techniques, increasing parallelism leads to better solutions faster, but only up to a certain point. At some degree of parallelism, the quality of solutions abruptly degrades to that of random sampling. This transition is sharp and displays many of the characteristics of thermodynamic phase transitions.

We compute critical exponents that characterize this transition using finite-size scaling, a technique borrowed from statistical physics, and demonstrate our results on two important problems: energy minimization on *NK* energy functions (*5*) and tour-length minimization for the traveling salesman problem (TSP) (*6*). The *NK* model, a generalization of spin-glass models (*7*), was chosen as one of the few general models capable of generating tunably difficult optimization tasks. The TSP is perhaps the most famous and well studied combinatorial optimization problem and is often used as a test-bed for new ideas. Both problems, although quite different, show remarkably similar behavior.

The *NK* model (*5*) defines a family of energy functions over a discrete search space. The search space consists of all possible configurations $\{s\} \equiv \{s_1, \ldots, s_N\}$ of N variables. If each of the variables can take any of A possible values, the search space is of size A^N. We confine our investigation to the case $A = 2$, call our variables spins, and let them take the values ± 1, because of their similarity with Ising spins, which arise in models of magnetism. Our local modifications for the *NK* model are single spin flips where $s_i \rightarrow -s_i$. Each spin s_i makes a contribution to the total energy dependent on its own state and the state of K other spins. These K other spins may be selected at random or according to some specified topology. We have considered both selection mechanisms, obtaining similar results in both cases. The total energy of a configuration $\{s\}$ is defined by

$$E\{s\} = \frac{1}{N}\sum_{i=1}^{N} E_i(s_i; s_{i_1}, \ldots, s_{i_K}) \qquad (1)$$

By analogy with spin glasses (*7*), the local energy contributions $E_i(s_i; s_{i_1}, \ldots, s_{i_K})$ for each of the 2^{K+1} local spin configurations, $\{s_i, s_{i_1}, \ldots, s_{i_K}\}$, are chosen at random from a uniform distribution over the range [0,1). However, by specializing E_i and the selection of the K neighbors, we can investigate specific optimization problems, for example, spin glasses, graph coloring, or number partitioning.

The number of other spins K that each spin interacts with varies from 0 to $N - 1$ and controls the ruggedness of the energy landscape. In the $K = 0$ limit, the spins are independent and there is a single minimum. As K increases, the number of conflicting constraints increases, leading to multiple local minima that can trap local search algorithms. At the $K = N - 1$ extreme, every spin affects every other spin, energies of adjacent configurations are uncorrelated, and there are exponentially many local minima. The $K \gg 1$ limit is analytically tractable and has been studied by many authors (*8*).

We focus on simulated annealing (*1*) as our representative local search algorithm. In simulated annealing, local modifications are accepted according to the Metropolis criterion (*9*), a method for simulating the evolution of a physical system to thermal equilibrium in a heat bath. A modification is applied and the resulting change in energy ΔE is computed. The modification is always accepted if $\Delta E \leq 0$ and is accepted with probability $p(\Delta E) = \exp(-\Delta E/T)$ if $\Delta E > 0$. The temperature parameter T controls the fraction of uphill moves that are accepted; at zero temperature, $T = 0$, only downhill moves are accepted. It is important to occasionally accept uphill moves (that is, $T > 0$) to prevent trapping on poor local minima where all local modifications raise the energy. Simulated annealing derives its name from the annealing, or gradual lowering, of the temperature parameter, so that uphill moves are accepted with decreasing frequency.

Beyond the obvious computational speedups of applying many local modifications simultaneously, parallel local search also makes it harder to get trapped in poor local optima. Even when only greedy moves are accepted ($T = 0$), allowing for simultaneous updates of the spins introduces outdated information and the possibility of "mistakes": if the energy of s_i depends on s_j and both spins simultaneously attempt to flip, s_i may conclude that $\Delta E < 0$ under the assumption that s_j did not flip; however, ΔE may actually be positive if s_j did flip. This similarity between parallelism and temperature is only superficial: Parallelism has qualitatively different effects on the behavior of optimization algorithms.

We parameterize the degree of parallelism by $0 < \tau \leq 1$, which denotes the probability that a spin attempts to flip. If there are N spins, then on average $N\tau$ of them are updating under the local operator at the same time. The $\tau = 1/N \rightarrow 0$ limit corresponds to a sequential algorithm, which at $T = 0$ will result in the system becoming trapped in (usually poor) local optima. Maximal parallelism is obtained for $\tau = 1$, when all spins update simultaneously. At this extreme, the system typically cannot converge because too many of the spins use outdated information.

The transition between these two ex-

W. G. Macready and S. A. Kauffman, Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.
A. G. Siapas, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: thanos@mit.edu
*To whom correspondence should be addressed.

tremes is surprisingly sharp. Figure 1 shows the asymptotic energy obtained under search at fixed temperature $T = 0$ (*10*). Qualitatively identical results are obtained under annealing when the temperature is cooled exponentially. An interesting feature of Fig. 1 is the initial improvement in optimizing performance with increasing τ. This improvement is most likely caused by the fact that parallel updating prevents premature trapping in poor local optima and guides the search toward low-energy regions of configuration space. Premature convergence is prevented because initially many spin flips decrease energy and are therefore accepted in parallel, resulting in large-scale configuration changes. Because each spin flip individually decreases energy, these large-scale changes move the system in a direction that, on average, appears to lead to low-energy regions of configuration space. Numerical experiments support this conjecture: Local energy minima around the configuration at which the parallel dynamics converges have energies that are lower than randomly sampled local minima by statistically significant amounts. As τ increases, more spins flip in parallel, more averaging is done, and coarser trends of the energy landscape are discerned. Therefore, the energy decrease with τ should be more significant for smoother landscapes (smaller K), where there are coarse features that can be exploited. This prediction is supported by the data in Fig. 1.

These observations point out two key differences between parallelism-induced noise and temperature-induced noise. (i) No phase transition occurs as temperature is varied, and (ii) in contrast to temperature-induced noise, which only depends on local energy differences, parallelism-induced noise also depends on long-range correlations in the energy landscape.

The abrupt degradation in performance of the parallelized algorithms is associated with an order-disorder transition in the update dynamics. Above the transition, where the energy remains high, spins continue to flip indefinitely. Below the transition, reaching low-energy configurations is accompanied by a freezing of the spins (at $T = 0$) to particular values. This observation motivates the definition of the following order parameter, which distinguishes the behavior on either side of the transition: Let p_i be the probability that at equilibrium the ith spin will flip if asked. We define the order parameter for the ith spin to be its entropy, $S_i = -p_i\log_2 p_i - (1 - p_i)\log_2(1 - p_i)$. This is 0 if the spin is frozen and 1 if the spin is flipping randomly. Images of the spatial distribution of S_i for a lattice of spins can be found in Fig. 2. The order parameter S for the entire system is defined as the average entropy per spin, $S = (1/N)\sum_{i=1}^{N} S_i$. For $K = 6$, the order parameter is close to 0 below the critical point τ_c and rises sharply towards 1 above τ_c (Fig. 3A).

The sharpness of the transition and the location of τ_c depend on the size N of the system (*11*). To analyze this dependence, we used finite-size scaling (*12*), a method from statistical physics in which the observation of how the critical point $\tau_c(N)$ changes with the size of the system gives direct evidence for critical behavior at the transition. As the size N increases, the transition sharpens and the transition point shifts according to

$$\tau_c(N) - \tau_c(\infty) \sim N^{-1/\nu} \qquad (2)$$

where ν is a critical exponent (Fig. 3B). The empirical observation behind this analysis is that sufficiently close to the critical point, systems of all sizes are indistinguishable except for an overall change of scale. If we define a rescaled parameter

$$y = N^{1/\nu}\frac{\tau - \tau_c(\infty)}{\tau_c(\infty)} \qquad (3)$$

the rescaled curves fall on a universal (N-independent) curve (Fig. 3C).

Analytical insight into the transition for the NK model can be obtained by making two approximations: all spins have the same energy E_t at time t (mean-field approximation) and energies induced by spin flips are chosen uniformly from the unit interval (annealed approximation). Then, $p_n(E) = \binom{K+1}{n}(\tau E)^n(1 - \tau E)^{K+1-n}$ is the probability that n spins flip given that the system is at energy E, and the time evolution of E is described by

$$E_{t+1} = p_0(E_t)E_t + p_1(E_t)\beta E_t + [1 - p_0(E_t) - p_1(E_t)]\Phi \qquad (4)$$

where $\beta E_t = (1/2)E_t$ is the expected energy that results from single flips that lower the energy, and Φ represents the "reset" energy that results when two or more interdependent spins flip simultaneously, modeled as $\Phi = (1 + \tau/5)/5$ (*13*). The fixed point of Eq. 4 is the asymptotic energy. Under the above simplifying assumptions, $E = 0$ is the only fixed point for low τ, and τ_c is the smallest τ for which a nonzero fixed point exists. By setting $E_{t+1} = E_t = E$ in Eq. 4 and expanding to second order in E, we find that $\tau_c(K) = 5(\sqrt{1 + 4/K} - 1)/2$ (*13*). A comparison of this prediction and the results from simulation is shown in Fig. 4. This result is consistent with the fact that the transition occurs only for $K > 4$ (*14*) because $\tau_c \geq 1$ for $K \leq 4$.

Thus far we have described an abrupt transition in a family of optimization problems as the degree of parallelism is varied. Qualitatively, at least, the forces driving the

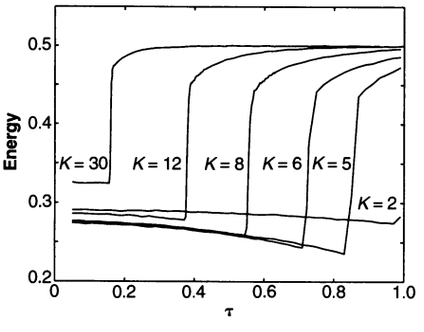

Fig. 1. Energy reached at the end of 500 generations versus τ for $N = 5000$, temperature $T = 0$, and several different K values. The search space is vast, of size $2^{5000} \approx 10^{1500}$. A generation is defined to be N update attempts, an average $N\tau$ of which are simultaneous. The average over 30 randomly generated NK landscapes is plotted for each τ value.

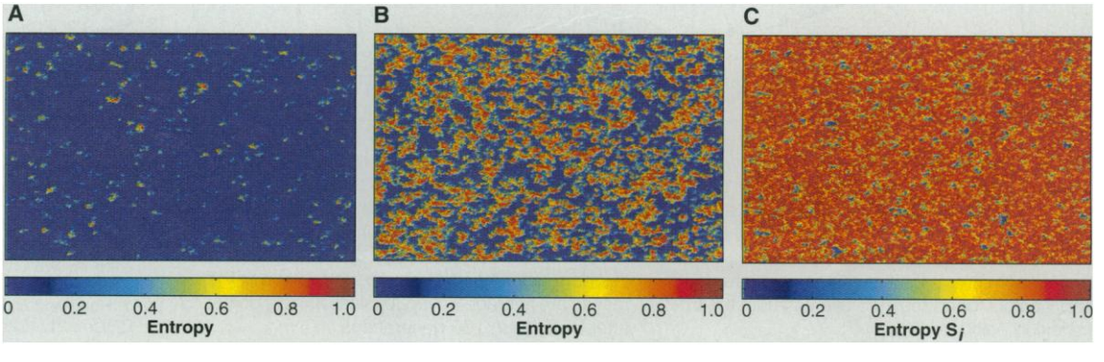

Fig. 2. Rendering of the entropy fields for a 800 × 800 lattice of spins at $T = 0$. The energy of each spin depends on the states of the nearest $K = 12$ spins in the lattice. The color of each spin represents the entropy computed over the last 40 of a total of 80 generations. The images correspond to (**A**) $\tau = 0.43 < \tau_c$, (**B**) $\tau = 0.47 \approx \tau_c$, and (**C**) $\tau = 0.51 > \tau_c$.

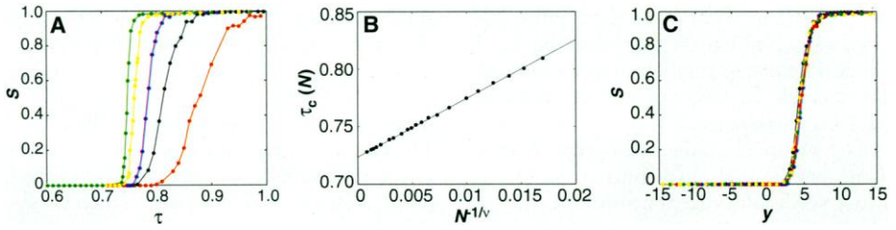

Fig. 3. (**A**) Order parameter of $K = 6$ energy functions with different N values ($N = 100$, red; $N = 200$, black; $N = 400$, blue; $N = 800$, yellow; and $N = 1600$, green). (**B**) $\tau_c(N)$ versus $N^{-1/\nu}$ for $K = 6$, where $\nu \approx 1.43 \pm 0.11$ and $\tau_c(\infty) \approx 0.720 \pm 0.005$, computed by a nonlinear least squares fit from a collection of N and $\tau_c(N)$ values. Here we define $\tau_c(N)$ to be the value of τ at which the entropy is 0.5. (**C**) The order parameter curves of (A) plotted with respect to $y = N^{1/\nu}(\tau - \tau_c)/\tau_c$ instead of τ. Within the error bars (not shown), all curves collapse onto a universal (N-independent) curve.

transition are not unique to the NK model. We provide further support for the generality of the observed phenomena by investigating a very different optimization task, the traveling salesman problem (TSP). Although part of its importance lies in its role as benchmark for new theory and optimization techniques, the TSP and its variants have many practical applications ranging from printed circuit-board design to x-ray crystallography to scheduling. The task is simple: find the shortest tour passing through a set of cities, visiting each city only once.

One of the most effective solution techniques for the TSP was found by Lin and Kernighan (*15*). Their method relies on a local operator, k-opt, to improve solutions. The k-opt operator removes k edges between cities and replaces them with k new edges such that the new edges still form a valid tour. At each improvement, the Lin-Kernighan heuristic intelligently looks for an improvement with any value k. Here we investigate the quality of solutions as a particular k-opt–like move is applied with increasing parallelism. The results do not depend on the exact form of the operator but only on the fact that it has a local interaction range. The operator we use is defined as follows: for city i we select $k - 1$ other cities at random, yielding a set $\{c_i, c_{i_1}, \ldots, c_{i_{k-1}}\}$ of k cities. This subset of cities is then cyclically permuted within the tour. For example, if the tour starting at city 1 and visiting cities 2, 3, . . . , 10, 1 in sequence is represented as {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} and the subset of cities considered for a 3-opt–like move is {3, 5, 7}, then the resulting tour after the move is {1, 2, 5, 4, 7, 6, 3, 8, 9, 10}. The results of applying this operator with increasing parallelism for a set of $N = 439$ cities and various interaction ranges, k, are presented in Fig. 5. A study of the order parameter for a wide variety of TSP instances reveals behavior analogous to that described for the NK model with k playing the role of K.

The phase transition that arises as local search algorithms are parallelized is driven by the synergistic and antagonistic effects of overlapping applications of local operators, and its existence places a sharp upper bound on the amount of useful parallelism in local search algorithms for combinatorial optimization. Techniques from statistical mechanics have been used to make the correspondence between the physics of phase transitions and the performance of parallelized optimizing systems more precise, demonstrating strong connections between these seemingly unrelated fields.

Phase transitions seen in the satisfiability of constraint satisfaction problems (*16*) have previously shown the importance of critical phenomena in artificial intelligence. The transition described here is quite distinct from this satisfiability transition but further demonstrates the importance of critical phenomena in optimization and artificial intelligence.

It is interesting to consider whether the critical behavior presented here arises in more general distributed systems of interacting agents. Any time the decision of one agent relies on information contained in the state of another agent, the possibility exists for an abrupt degradation in performance as more agents act in parallel and the amount of "stale" information that exists because of nonzero information-propagation delays increases. A similar transition seen in the transmission of packets in computer networks as the traffic density rises (*17*) provides support for the generality of these phenomena. Such transitions may also be lurking in some human organizations.

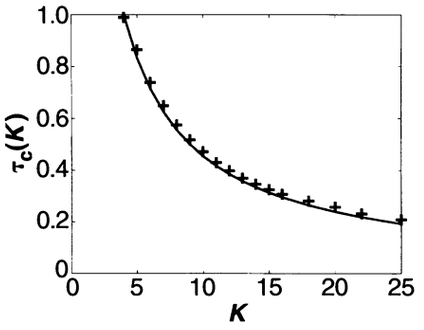

Fig. 4. Comparison of the prediction $\tau_c(K) = 5(\sqrt{1 + 4/K} - 1)/2$ (solid curve) with estimates of τ_c at which the entropy is 0.5 (data points represented by +s). Here $N = 1000$ and $K = 1, \ldots, 25$, and parallelized greedy search is used ($T = 0$). Because N is large, the transition curves are sharp and τ points for entropy 0.5 overestimate the true transition point very little.

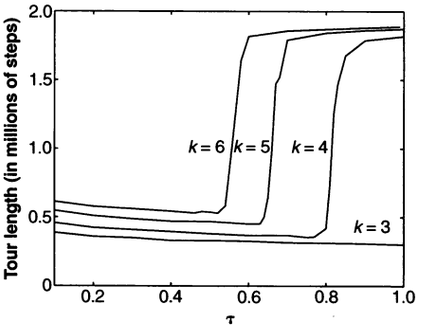

Fig. 5. Expected tour lengths under $k = 3$, 4, 5, and 6 local moves as a function of τ for a set of $N = 439$ cities called pr439.tsp. This tour is supplied in the TSPLIB package, which is available by anonymous ftp through elib.zib-berlin.de. The results are averaged over 30 initial starting points after having run 100 generations.

REFERENCES AND NOTES

1. S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, *Science* **220**, 671 (1983).
2. S. Forrest, *ibid.* **261**, 872 (1993).
3. F. Glover, *ORSA J. Comput.* **1**, 190 (1989); *ibid.* **2**, 4 (1990); D. Cvijović and J. Klinowski, *Science* **267**, 664 (1995).
4. R. Azencott, Ed., *Simulated Annealing Parallelization Techniques* (Wiley, New York, 1992).
5. S. A. Kauffman, *The Origins of Order* (Oxford Univ. Press, Oxford, 1993).
6. G. Reinelt, *The Traveling Salesman: Computational Solutions for TSP Applications* (Springer-Verlag, Berlin, 1994).
7. K. H. Fischer and J. A. Hertz, *Spin Glasses* (Cambridge Univ. Press, Cambridge, 1991).
8. B. Derrida, *Phys. Rev. B* **24**, 2613 (1981); C. A. Macken and A. S. Perelson, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6191 (1989); E. D. Weinberger, *Phys. Rev. A* **44**, 6399 (1991).
9. N. A. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. J. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
10. This transition persists for small, nonzero, fixed temperatures and broadens at larger temperatures.
11. The position of the transition also depends on the length of time that we allow the algorithm to run. In all of our numerical experiments, we waited sufficiently long that further increases in running time did not noticeably affect the transition.
12. R. J. Creswick *et al.*, *Introduction to Renormalization Group Methods in Physics* (Wiley, New York, 1992); J. J. Binnet *et al.*, *The Theory of Critical Phenomena* (Oxford Univ. Press, Oxford, 1992).
13. A. G. Siapas and W. G. Macready, in preparation.
14. Without interaction ($K = 0$), there will be no transition; this effect persists up to some nonzero degree of interaction (Fig. 1). It requires a sufficient amount of conflict between variables to overcompensate for the downward trend in energy obtained by accepting (in parallel) only spin flips that individually lower the energy.
15. S. Lin and B. W. Kernighan, *Oper. Res.* **21**, 498 (1973).
16. S. Kirkpatrick and B. Selman, *Science* **264**, 1297 (1994); P. Cheeseman, B. Kanefsky, W. M Taylor, in *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, (Morgan Kaufmann, San Mateo, CA, 1991), vol. 1, pp. 331–337.
17. T. Hogg and B. Huberman, *IEEE Trans. Syst. Man Cybern.* **21**, 1325 (1991).

18. We thank D. Stein for helpful discussions, and H. Abelson, G. J. Sussman, J. Tsitsiklis, E. Hung, N. Berker, E. Todorov, J. Tenenbaum, A. Khwaja, P. Beebee, and the reviewers for useful suggestions. This work was partially supported by Advanced Research Projects Agency grant ARPA ONR-N00014-92-J-4097 and National Science Foundation grant NSF MIP-9001651 (A.G.S.), and by Department of Energy grant DOE DE-FG03-93-ER61639, NSF BIR-9218746, and National Institutes of Health grant 5-R01 GM49619 (W.G.M. and S.A.K.).

1 June 1995; accepted 3 November 1995

Quantifying Hydrogen Bond Cooperativity in Water: VRT Spectroscopy of the Water Tetramer

J. D. Cruzan, L. B. Braly, Kun Liu, M. G. Brown, J. G. Loeser, R. J . Saykally*

Measurement of the far-infrared vibration-rotation tunneling spectrum of the perdeuterated water tetramer is described. Precisely determined rotational constants and relative intensity measurements indicate a cyclic quasi-planar minimum energy structure, which is in agreement with recent ab initio calculations. The O–O separation deduced from the data indicates a rapid exponential convergence to the ordered bulk value with increasing cluster size. Observed quantum tunneling splittings are interpreted in terms of hydrogen bond rearrangements connecting two degenerate structures.

The quest for a genuine molecular description of bulk-phase water has motivated two lines of research. The first entails a convergence of theory and experiment in determining an accurate six-dimensional intermolecular pair potential hypersurface (IPS). As a central element of this effort, several of the intermolecular vibrational modes of the water dimer have been characterized by far-infrared vibration-rotation tunneling (FIR-VRT) spectroscopy (*1*). The second direction involves investigation of the structures and dynamics of larger water clusters, with the principal goal of quantifying the nature of cooperative (many-body) effects in hydrogen bonding (*2*). Detailed observation of the water trimer with tunable FIR-VRT spectroscopy confirmed that the equilibrium structure of the trimer is cyclic and quasi-planar, as predicted by many theoretical calculations (*3*, *4*). Subsequent spectroscopic observations (*5*, *6*) and several theoretical studies (*7–9*) have elaborated on this work.

Here we report measurement of the FIR-VRT spectrum of the fully deuterated water tetramer. The structure of this molecule, as deduced from its spectrum, is in agreement with most theoretical predictions (*4*, *10–12*), which indicate that the minimum energy structure should be cyclic and quasi-planar (Fig. 1A). From these new data, we can begin to quantify the four-body terms in the interaction potential of bulk water.

The tunable FIR spectrometers used in this work have been described previously (*12–14*) (Fig. 2). Forty VRT transitions were measured to ±1.5 MHz in the spectral region near 68.0 wave numbers (cm^{-1}) and assigned to a parallel band of a symmetric rotor molecule. Table 1 lists the rotational constants obtained by nonlinear regression from the energy expression

$$E(J,K) = BJ(J+1) + (C-B)K^2 - D_J J^2(J+1)^2 + D_{JK}J(J+1)K^2 \quad (1)$$

where B and C are the symmetric top ($A = B$) molecular rotational constants, D_J and D_{JK} are centrifugal distortion parameters, and J and K are the assigned rotational quantum numbers. Each vibration-rotation transition was observed as a doublet of lines with approximately equal intensity and a constant (no change with J or K) average spacing of 5.6 ± 0.1 MHz (Fig. 2). Each datum in the fit is the numerical average of the two doublet frequencies.

Because a distribution of cluster sizes and species is present in a supersonic expansion, several tests were undertaken to identify the carrier of the spectra. Identical spectra were observed in expansions of Ar, Ne, and He; hence clusters containing carrier gas atoms were ruled out. As outlined in (*13*), the number of D_2O molecules in the cluster was determined by measurement of the intensity of several spectral features as a function of the mole fraction of D_2O in an H_2O-D_2O mixture. The amount of D_2O in the gas mixture can be varied while the total number of water molecules available for clustering is conserved. In this model, isotopic effects on the rates of cluster formation are assumed to be negligible. Five isotopic mixtures were tested and the resulting data were fit to the relation $\ln(I) = 2n \cdot \ln(\chi_{D_2O})$, where I is the relative signal intensity, χ_{D_2O} is the mole fraction of D_2O, and n is the number of D_2O molecules in the cluster. The results (Fig. 3) yield $n = 4.14 \pm 0.22$ at the 95% confidence level. Because no signal increase was observed with the use of any of the isotopic mixtures, clusters *containing mixed water* isotopes were ruled out as well. Additionally, the rotational constants, determined to a precision of about one part in 10^4, cannot be accounted for by consideration either of pure D_2O clusters containing more or less than four monomers or of other clusters containing any obvious contaminants.

Parallel transitions ($\Delta K = 0$) do not allow absolute determination of the C rotational constant in either state from the spectral line positions. Only the relative difference between C constants in the upper and lower states could be determined from these data, which precludes a rigorous rotational structural determination. Relative transition intensities, however, can yield an approximate value for the C rotational constant, because they arise from the absolute energy level positions in the lower vibrational state, which depend on that parameter. Comparison of careful relative intensity measurements in the compact Q-branch ($\Delta J = 0$) region with simulated intensity profiles as the C constant was varied between limits that would yield spherical ($A = B = C$) and planar oblate ($A = B = 2C$) structures (Fig. 2B) showed that the most consistent value was for a near-planar structure ($C = 1500 \pm 250$ MHz).

The consensus of most ab initio quantum chemistry studies is that the minimum energy structure of the water tetramer is similar to that shown in Fig. 1A, which is best described by the S_4 point symmetry group. Recent high-level ab initio calculations by Xantheas (*9*) have yielded the rotational constants $A = B = 3149$ MHz and $C = 1622$ MHz for the S_4 structure, with an O–O separation of 2.74 Å. Assuming an S_4 structure in the present case and scaling for the experimentally determined rotational constants yield approximate average O–O separations of 2.78 and 2.79 Å in the lower and upper states, respectively. This result is consistent with the observed trend in O–O separations of 2.976, 2.90, and 2.76 Å in the water dimer (*15*), trimer (*8*), and pentamer (*16*) and indicates that this property is converging rapidly to the ordered bulk value with cluster size. Figure 4 shows that the experimental O–O separations follow the trends predicted by Xantheas's Hartree-Fock (HF) and second-order Möller-Plesset (MP2) calculations (*9*).

Spectral splittings, such as the doublets observed in this work, have been observed in the spectra of many small clusters, in-

Department of Chemistry, University of California at Berkeley, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed.