## Kinetic Transition Networks for the Thomson Problem and Smale's Seventh Problem

Dhagash Mehta,<sup>1,2,3,\*</sup> Jianxu Chen,<sup>4,†</sup> Danny Z. Chen,<sup>4,‡</sup> Halim Kusumaatmaja,<sup>5,§</sup> and David J. Wales<sup>6,||</sup>

<sup>1</sup>Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, Indiana 46556, USA

<sup>2</sup>Centre for the Subatomic Structure of Matter, Department of Physics, School of Physical Sciences, University of Adelaide,

Adelaide, South Australia 5005, Australia

<sup>3</sup>Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA

<sup>4</sup>Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA

<sup>5</sup>Department of Physics, Durham University, South Road, Durham DH1 3LE, United Kingdom <sup>6</sup>University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, United Kingdom

(Received 10 December 2015; published 6 July 2016)

The Thomson problem, arrangement of identical charges on the surface of a sphere, has found many applications in physics, chemistry and biology. Here, we show that the energy landscape of the Thomson problem for N particles with N = 132, 135, 138, 141, 144, 147, and 150 is single funneled, characteristic of a structure-seeking organization where the global minimum is easily accessible. Algorithmically, constructing starting points close to the global minimum of such a potential with spherical constraints is one of Smale's 18 unsolved problems in mathematics for the 21st century because it is important in the solution of univariate and bivariate random polynomial equations. By analyzing the kinetic transition networks, we show that a randomly chosen minimum is, in fact, always "close" to the global minimum in terms of the number of transition states that separate them, a characteristic of small world networks.

DOI: 10.1103/PhysRevLett.117.028301

Introduction.—The Thomson problem is an important model in both physics and chemistry and is easily stated: find the minimum energy of a system composed of N identical charges that are constrained to move on the surface of a sphere of unit radius, i.e., minimize the potential energy function

$$V_{\rm Th}(\mathbf{r}) = \sum_{1 \le i < j \le N} \frac{1}{r_{ij}},\tag{1}$$

where  $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$ , subject to spherical constraints  $x_i^2 + y_i^2 + z_i^2 = 1$  for all i = 1, ..., N.

The model was proposed by J. J. Thomson as a representation of atomic structure [1], although it was soon abandoned in this context. However, it has subsequently found many applications in physics, chemistry, and biophysics because it captures the competition between local order for neighboring particles and long-range constraints due to the curvature and geometry. In particular, it is not generally possible to arrange every particle in an identical environment. The resulting model can provide insight into the forces governing far more complex systems, such as the arrangement of proteins in shells of spherical viruses [2–5], fullerene patterns for carbon clusters [6], the surface ordering of liquid metal drops confined in Pauli traps [7], and the behavior of colloidal particles trapped at a fluid-fluid interface [8–15].

In the mathematics and computer science communities, the Thomson problem has gained special attention because it appears in the 7th problem in Steven Smale's list of eighteen unsolved problems for the 21st century [16,17]. The problem, as posed by Smale [18], requires us to algorithmically construct a collection of starting points [21], say  $\mathbf{r}^*$ , so that for a given number of particles N,  $V_{\text{Th}}(\mathbf{r}^*) - V_{\text{Th}}(\mathbf{r}_{\text{gm}}) \leq \text{const} \times \log N$ , for  $N \geq 2$ . Here,  $\mathbf{r}_{\text{gm}}$  is the global minimum for the given N. Constructing such points corresponds to finding a good starting system of polynomial equations to locate all the complex solutions for sets of polynomial equations and to realize the fundamental theorem of algebra [22,23].

Given its broad relevance, the global minimum of the Thomson problem has been studied extensively [24–37]. Analytical solutions are known for N = 2-6 and 12 [38,39]. For N = 4, 6, and 12, they correspond to Platonic solids. The global minimum structures for larger *N* have been tackled computationally [32,33,40–48]. Most of these studies, however, only focus on the global minima. In the present Letter, we search extensively for (most, if not all) the minima and transition states of the Thomson problem at some selected sizes up to N = 150 and study the networks defined by these minima and the transition states that connect them. Using disconnectivity graphs [49,50], we find that the Thomson problem exhibits a single funnel for all N considered. We also show that the networks exhibit typical characteristics of a small world network. Hence, in the context of Smale's 7th problem, all minima are only a few steps from the global minimum.

Potential energy landscapes of the Thomson problem.— We begin by collecting some basic definitions: a local minimum of a potential is a point in the configuration space where the gradient vanishes and at which the Hessian matrix has no negative eigenvalues. The minimum at which the potential attains the lowest value is the global minimum. A transition state [51] is a configuration where the gradient vanishes, and exactly one eigenvalue of the Hessian matrix is negative. Transition states define connections between pairs of minima via steepest-descent paths. Hence, we can construct kinetic transition networks [52–54] of local minima and the transition states that link them for a given potential.

To perform an extensive search for local minima and transition states of the Thomson problem, we employed the GMIN [55] and OPTIM programs [56]. To identify likely global minima, we applied basin-hopping global optimization [57–60]. In this method, random geometrical perturbations are followed by energy minimization, and moves are accepted or rejected based upon the energy differences between local minima. This procedure transforms the energy landscape system into the set of catchment basins for the local minima. For all the minimizations in the present Letter, a modified version of the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm [61,62] was used. This scheme has proved to be the most efficient in recent benchmarks [63].

We used a combination of the doubly-nudged elastic band (DNEB) and hybrid eigenvector-following techniques to find the transition states [64]. In DNEB, a series of images interpolate between the two end points, and the total energy is minimized subject to spring constraints between adjacent images. Maxima in the DNEB path are then adopted as transition state candidates, which we refine to high precision using hybrid eigenvector-following [65–67]. Here we use a Rayleigh-Ritz approach [60] to compute the smallest nonzero eigenvalue, and take an uphill step along the corresponding eigendirection. Then we minimize in the tangent space for a limited number of steps, so that the gradient does not develop a significant component in the uphill direction. We project out all components corresponding to zero Hessian eigenvalues. For the Thomson problem, these are the three modes associated with overall rotations around the x, y, and z axes. Using spherical polar coordinates  $(\phi, \theta)$ , the corresponding eigenvectors are  $\hat{e}_x = (\cos\theta\cos\phi/\sin\theta, \sin\phi), \hat{e}_y = (-\cos\theta\sin\phi/\sin\theta, \cos\phi)$ and  $\hat{e}_{z} = (1, 0)$ .

For each transition state, we applied small displacements in the two downhill directions and minimized the energy to identify minimum-transition state-minimum triplets. In some cases, new local minima may be identified in this procedure. By iterating the process, we systematically build a connected database of stationary points.

Our results for the number of minima and transition states at N = 132, 135, 138, 141, 144, 147, and 150 are shown in Fig. 1. In all cases [68], we have obtained more minima than previous results given in the literature [27,69]. It is known that the number of local minima for a molecular system usually increases exponentially with system size [27,70,71]. This scaling applies here, and locating likely



FIG. 1. "TS" refers to "Energetically Distinct Transition Sates," "EH" refers to the estimate of the number of minima given in [27], "CGS1" is a recent [69] fit to the number of minima found in previous calculations, and "CGS2" is an estimate [69] for the number of minima (we use the maximum out of those suggested in [69]). The lines connecting data points are a guide to the eve.

candidates for the global minimum becomes increasingly difficult for larger systems. However, the rate of increase is relatively slow for the Thomson problem, because the Coulomb potential is long-ranged [72–74].

Disconnectivity graphs provide a powerful way to visualize the organization of minima and transition states [49,50]. A disconnectivity graph is a tree graph where the vertical axis corresponds to the potential energy. Each line terminates at the energy of a local minimum, and the minima are joined at the lowest energy where they can interconvert for regularly spaced energy thresholds. These connections are defined by the highest transition state on the lowest energy path between each pair of minima.

The disconnectivity graph for the databases considered here is shown in Fig. 2 for N = 147. It shows a typical structure-seeking "palm tree" organization [50], where there is only a single funnel in the potential energy landscape, and the energy barriers separating local energy minima and the global minimum are small. The same qualitative features are observed for all particle numbers we have considered, up to N = 150. This organization is associated with efficient relaxation to the global minimum.

*Network properties.*—For each N, we construct a network G in which each node corresponds to a unique minimum energy structure, where permutation-inversion isomers are lumped together. Hence, there are  $\mathcal{N}_{min}$  nodes. Two nodes are connected by an edge if a transition state exists between the corresponding minima. For each N, we have obtained  $\mathcal{N}_{ts}$  energetically distinct transition states. Here, we need only consider whether two minima are connected or not, and more than one distinct transition state between a pair of minima only counts as a single edge. We call  $\mathcal{N}_{\text{edge}}$  the number of edges for the given network, with  $\mathcal{N}_{ts} \geq \mathcal{N}_{edge}$ . As in previous results for Lennard-Jones clusters [75-77], our networks are undirected and unweighted graphs and, hence, agnostic about all other information, such as barrier heights or transition rates between the minima.



FIG. 2. (a) The disconnectivity graph for N = 147. In all cases considered here, the disconnectivity graphs show a typical structure-seeking "palm tree" organization. In panel (b), the insets show the structures of the lowest five minimum energy configurations. The pentagons, hexagons, and heptagons indicate that the charges in the center of the polygons have five, six, and seven neighbors, respectively.

To further describe our findings, we first define two quantities. The characteristic path length (L) is the average of the shortest path between each pair of nodes in G. We apply an all-pairs shortest path algorithm [78] to the graph G to compute the distance between every two nodes. L is

the arithmetic mean of the distances between all pairs of nodes with a finite distance. The local clustering coefficient  $C_i$  of the *i*th node is defined as the fraction of pairs of neighboring nodes that are connected over all pairs of neighbors of the node. The global clustering coefficient *C* is the average of  $C_i$  for  $i = 1, ..., N_{min}$ .

A graph is a small-world network if the characteristic path length is similar to, and the clustering coefficient is much higher than, the random graph with the same number of nodes and edges [79]. Small-world properties have been observed in a number of networks of dynamic systems. Our results for the networks of minima of the Thomson problem for various values of N are reported in Table I. The characteristic path length and the clustering coefficient of the corresponding random graph are denoted as  $L_0$  and  $C_0$ , respectively. It is evident that L is as small as  $L_0$ , while C is much larger than  $C_0$ , and hence, the small-world characteristics of the networks are established. The network diameter values, defined in terms of the longest shortest path length, are 5 (N = 132, 135, 138, 141, 144), 6 (N = 147), and 7 (N = 150).

Another important network property is the betweenness centrality, which is defined for each node as the number of shortest paths from all nodes to all others that pass through the node under consideration. Betweenness centrality determines how central the node is. In Fig. 3, we plot the betweenness centrality of all the nodes of our networks for N = 138. Our results reveal that the global minimum of the Thomson problem is the most central in the network. The number of connections (the node degree) of minima as a function of energy is plotted in Fig. 4. The figure clearly shows that low-lying minima are highly connected, making them hubs, unlike the higher energy minima.

Discussion and conclusion.-In this Letter, we have characterized the potential energy landscape of the Thomson problem for selected sizes up to N = 150. We confirmed that both the number of minima and transition states grow exponentially with N, albeit with a small exponential factor due to the long range nature of the Coulomb interactions. This exponential increase makes the searches for the global minima of the Thomson problem progressively more difficult. However, by analyzing the disconnectivity graphs, we find that for the sizes investigated here, the landscapes exhibit clear structure-seeking organization. The global minima for systems characterized by such funneled potential energy landscapes can usually be located easily, even when the total number of minima is large. The asymmetry of the energy barriers makes it easy to escape from high to low energy minima, while the reverse transitions are significantly slower.

We have analyzed the networks consisting of local minima and the connections between them defined by transition states and find that they exhibit small world properties [79]. Our results provide further evidence to support the conjecture that the small-world phenomenon

here $L_0$ and $C_0$ are the characteristic pair length and the clustering coefficient of the corresponding random graph.							
N	$\mathcal{N}_{\min}$	${\cal N}_{ m ts}$	$\mathcal{N}_{\mathrm{edge}}$	L	$L_0$	С	$C_0$
132	1183	28 284	8700	2.3003	2.8959	0.6032	0.0124
135	1585	51 832	11 285	2.2368	3.0391	0.6932	0.009
138	3226	88 999	20 1 50	2.1791	3.4865	0.6312	0.0039
141	4165	100 085	38 210	2.5684	3.1487	0.5871	0.0044
144	4534	136 519	33 113	2.3369	3.4445	0.7142	0.0032
147	6644	151 299	39 241	2.4314	3.8247	0.6375	0.0018
150	9774	178 728	87 203	2.7601	3.5163	0.4312	0.0018

TABLE I. Analysis of networks for different values of N.  $N_{min}$ ,  $N_{ts}$ , and  $N_{edge}$  are the number of minima (nodes), energetically distinct transition states, and edges, respectively. L and C are the characteristic path length and the clustering coefficient of each 4 1 4 14 17 1 · · · ·

might be another generic feature of landscapes with the structure-seeking palm tree organization [80]. Moreover, we found that the low-lying minima are generally significantly more connected than those at higher energy, suggesting scale-free properties. Further statistical tests [81] are required to confirm this behavior. We note that similar kinetic hubs have been identified for biomolecules in previous publications [80,82,83]. It is important to note that these networks are static. Hence, the scale-free phenomenon, if confirmed, requires explanation beyond the usual preferential attachment schemes [84].

Our results are relevant for addressing Smale's 7th problem. For the sizes considered here, the global minimum is, on average, only a few transition states  $[O(\log N_{\min})]$  away from any random starting point.  $\mathcal{N}_{\min}$  grows exponentially with N, and the increase in the average number of connections is linear. Since the diameter of the networks is typically five or six (and at most seven), the global minimum is never further than a few steps away, even from the highest-lying minimum. Moreover, the betweenness centrality is largest, by orders of magnitude, for the global minimum, so the majority of shortest paths between all pairs of minima pass through the global minimum; i.e., the global minimum is the central node of these networks. Hence, finding the global minimum and exploiting the funneled or small-world structure



FIG. 3. Betweenness centrality of the individual minima for N = 138.

of the landscape, is relatively straightforward from a numerical optimization point of view. Interestingly, though our results suggest that finding the global minimum of the Thomson problem may be relatively easy, finding an answer to the Smale problem is nondeterministic polynomial-time hard [85]. Constructing and analyzing networks of minima for the logarithmic version of the potential [17] may provide more concrete details on the mathematics behind Smale's 7th problem.

The present results agree with the observation [75] that funneled energy landscapes display small world characteristics. In the future, we will explore whether these two features are generally correlated, or if a counterexample can be found. Another avenue for future work is to extend our analysis to larger particle numbers. Our previous results for the Thomson problem [33] found that global minima for N > 400 start to display alternative defect motifs. We expect the potential energy landscape to display multiple funnels in this regime, and it will be interesting to see whether the small world phenomenon found here will be preserved.

In the future, we will further analyze the network properties of this model by including weights and directions for the edges, depending on the barrier heights and

Number of Connections  $10^{2}$ 10 10 8627.6 8627.5 8627.7 8627.8 8627.9 8628 8628.1 Energy

FIG. 4. The number of connections for local minima is plotted as a function of potential energy for N = 138. The line is the binned average.

10<sup>4</sup>

10

kinetic transition rates. We also plan to develop more specific algorithms to locate local and global minima by exploiting small-world properties [86–88]. Analyzing the appropriately weighted and directed networks of free energy minima [89] and transition states may provide additional insight into the Thomson problem. Furthermore, investigating network properties of higher index saddles may also provide insights for developing novel optimization algorithms [90].

D. M. was supported by Australian Research Council DECRA Fellowship No. DE140100867. D. Z. C. was supported in part by the NSF under Grant No. CCF-1217906. We thank Carlos Beltrán and Edward Saff for their helpful remarks and the authors of Ref. [69] for clarifying their results.

<sup>\*</sup>dmehta@nd.edu <sup>†</sup>jchen16@nd.edu <sup>‡</sup>dchen@nd.edu <sup>§</sup>halim.kusumaatmaja@durham.ac.uk <sup>|</sup>dw34@cam.ac.uk

- J. J. Thomson, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 7, 237 (1904).
- [2] D. L. Caspar and A. Klug, in *Cold Spring Harbor Symposia* on *Quantitative Biology* (Cold Spring Harbor Laboratory Press, New York, 1962), Vol. 27, pp. 1–24.
- [3] C. J. Marzec and L. A. Day, Biophys. J. 65, 2559 (1993).
- [4] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, Phys. Rev. Lett. 90, 248101 (2003).
- [5] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, Proc. Natl. Acad. Sci. U.S.A. 101, 15556 (2004).
- [6] H. W. Kroto et al., Nature 318, 162 (1985).
- [7] E. J. Davis, Aerosol Sci. Technol. 26, 212 (1997).
- [8] A. R. Bausch et al., Science 299, 1716 (2003).
- [9] P. Lipowsky, M. J. Bowick, J. H. Meinke, D. R. Nelson, and A. R. Bausch, Nat. Mater. 4, 407 (2005).
- [10] T. Einert, P. Lipowsky, J. Schilling, M. J. Bowick, and A. R. Bausch, Langmuir 21, 12076 (2005).
- [11] H. Kusumaatmaja and D. J. Wales, Phys. Rev. Lett. 110, 165502 (2013).
- [12] G. M. J. Paulose, D. R. Nelson, and V. N. Manoharan, Science 343, 634 (2014).
- [13] W. T. M. Irvine, V. Vitelli, and P. M. Chaikin, Nature (London) 468, 947 (2010).
- [14] W. T. M. Irvine, M. J. Bowick, and P. M. Chaikin, Nat. Mater. 11, 948 (2012).
- [15] E. Bendito, M. J. Bowick, A. Medina, and Z. Yao, Phys. Rev. E 88, 012405 (2013).
- [16] S. Smale, Math. Intell. 20, 7 (1998).
- [17] S. Smale, Fields Institute, 1997 (unpublished).
- [18] Smale's 7th problem, rather, considers a general form of the potential,  $r^{\alpha}$ , with the spherical constraints. The special values  $\alpha = 0$  and 1 correspond to a logarithmic potential and the Thomson problem, respectively. In Ref. [17],  $\alpha = 0$  is considered as the "main" problem (and the Thomson problem listed as a special case of the general problem)

and more mathematicians are focused on this version of the problem.  $\alpha = 1$  is a popular model in chemistry and physics. In the asymptotic limit, both versions of the problems are similar in the following sense: a solution to the Thomson problem is known to have an asymptotically optimal logarithmic potential. This statement is proven both ways in Refs. [19,20]. In any case, both versions of the problem share the same importance and qualitative features.

- [19] P. Leopardi, Adv. Comput. Math., 39, 27 (2013).
- [20] C. Beltrán, Found. Comput. Math. 15, 125 (2015).
- [21] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation* (Springer, New York, 2012).
- [22] M. Shub and S. Smale, J. Complexity 9, 4 (1993).
- [23] D. Armentano, C. Beltrán, and M. Shub, Trans. Am. Math. Soc. 363, 2955 (2011).
- [24] L. T. Wille, Nature (London) 324, 46 (1986).
- [25] T. Erber and G. M. Hockney, J. Phys. A 24, L1369 (1991).
- [26] E. L. Altschuler, T. J. Williams, E. R. Ratner, F. Dowla, and F. Wooten, Phys. Rev. Lett. 72, 2671 (1994).
- [27] T. Erber and G. M. Hockney, Phys. Rev. Lett. **74**, 1482 (1995).
- [28] E. L. Altschuler, T. J. Williams, E. R. Ratner, R. Tipton, R. Stong, F. Dowla, and F. Wooten, Phys. Rev. Lett. 78, 2681 (1997).
- [29] T. Erber and G. M. Hockney, Adv. Chem. Phys. **98**, 495 (1997).
- [30] J. R. Morris, D. M. Deaven, and K. M. Ho, Phys. Rev. B 53, R1740 (1996).
- [31] E. L. Altschuler and A. Pérez-Garrido, Phys. Rev. E 71, 047703 (2005).
- [32] E. L. Altschuler and A. Pérez-Garrido, Phys. Rev. E 73, 036108 (2006).
- [33] D. J. Wales and S. Ulker, Phys. Rev. B 74, 212101 (2006).
- [34] D. J. Wales, H. McKay, and E. L. Altschuler, Phys. Rev. B 79, 224115 (2009).
- [35] J. S. Brauchart, D. P. Hardin, and E. B. Saff, Bull. Lond. Math. Soc. 41, 621 (2009).
- [36] J. Brauchart, D. Hardin, and E. Saff, Contemp. Math. 578, 31 (2012).
- [37] M. Bowick, C. Cecka, and A. Middleton, Thomson problem @ s.u., http://physics.syr.edu/condensedmatter/ thomson/thomson.htm.
- [38] N. N. Andreev, East J. Approx. 2, 459 (1996).
- [39] R.E. Schwartz, arXiv:1001.3702.
- [40] E. B. Saff and A. B. Kuijlaars, Math. Intell. 19, 5 (1997).
- [41] A. Kuijlaars and E. Saff, Trans. Am. Math. Soc. 350, 523 (1998).
- [42] A. Pérez-Garrido and M. A. Moore, Phys. Rev. B 60, 15628 (1999).
- [43] A. M. Livshits and Y. E. Lozovik, Chem. Phys. Lett. 314, 577 (1999).
- [44] M. J. Bowick, D. R. Nelson, and A. Travesset, Phys. Rev. B 62, 8738 (2000).
- [45] M. Bowick, A. Cacciuto, D. R. Nelson, and A. Travesset, Phys. Rev. Lett. 89, 185502 (2002).
- [46] M. J. Bowick, A. Cacciuto, D. R. Nelson, and A. Travesset, Phys. Rev. B 73, 024115 (2006).
- [47] L. Giomi and M. J. Bowick, Phys. Rev. E 78, 010601 (2008).
- [48] T. LaFave, Jr., J. Electrost. 72, 39 (2014).

- [49] O. M. Becker and M. Karplus, J. Chem. Phys. 106, 1495 (1997).
- [50] D. J. Wales, M. A. Miller, and T. R. Walsh, Nature (London) 394, 758 (1998).
- [51] J. N. Murrell and K. J. Laidler, Trans. Faraday Soc. 64, 371 (1968).
- [52] F. Noé and S. Fischer, Curr. Opin. Struct. Biol. 18, 154 (2008).
- [53] D. Prada-Gracia, J. Gómez-Gardenes, P. Echenique, and F. Fernando, PLoS Comput. Biol. 5, e1000415 (2009).
- [54] D. J. Wales, Curr. Opin. Struct. Biol. 20, 3 (2010).
- [55] D. J. Wales, GMIN: A Program for Basin-Hopping Global Optimization, Basin-Sampling, and Parallel Tempering, http://www-wales.ch.cam.ac.uk/software.html.
- [56] D. J. Wales, OPTIM: A Program for Geometry Optimization and Pathway Calculations, http://www-wales.ch.cam.ac.uk/ software.html.
- [57] Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. 84, 6611 (1987).
- [58] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A 101, 5111 (1997).
- [59] D. J. Wales and H. A. Scheraga, Science 285, 1368 (1999).
- [60] D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, England, 2003).
- [61] J. Nocedal, Math. Comput. 35, 773 (1980).
- [62] D. Liu and J. Nocedal, Math. Program. 45, 503 (1989).
- [63] D. Asenjo, J. D. Stevenson, D. J. Wales, and D. Frenkel, J. Phys. Chem. B 117, 12717 (2013).
- [64] S. A. Trygubenko and D. J. Wales, J. Chem. Phys. 120, 2082 (2004).
- [65] G. Henkelman, B. P. Uberuaga, and H. Jónsson, J. Chem. Phys. 113, 9901 (2000).
- [66] L. J. Munro and D. J. Wales, Phys. Rev. B 59, 3969 (1999).
- [67] Y. Kumeda, L. J. Munro, and D. J. Wales, Chem. Phys. Lett. 341, 185 (2001).
- [68] We compare our results with the estimate on the number of minima,  $0.382 \exp(0.0497N)$ , as given in [27], the number of minima from a fit to the number of minima found,  $(0.31701 \pm 0.1) \exp[(0.0518 \pm 0.0012)N]$  and estimated number of minima  $\exp(-3.97635 \pm 0.1992)$

 $\exp[(0.078\,9298\pm 0.001\,76)N]$  (choosing the "+" sign for both "±" in our comparison), as given in [69].

- [69] M. Calef, W. Griffiths, and A. Schulz, J. Stat. Phys. 160, 239 (2015).
- [70] F. H. Stillinger and T. A. Weber, Science 225, 983 (1984).
- [71] D. J. Wales and J. P. K. Doye, J. Chem. Phys. 119, 12409 (2003).
- [72] J. P. K. Doye, D. J. Wales, and R. S. Berry, J. Chem. Phys. 103, 4234 (1995).
- [73] J. P. K. Doye and D. J. Wales, Science 271, 484 (1996).
- [74] J. P. K. Doye and D. J. Wales, J. Chem. Soc., Faraday Trans. 93, 4233 (1997).
- [75] J. P. K. Doye, Phys. Rev. Lett. 88, 238701 (2002).
- [76] J. P. Doye and C. P. Massen, J. Chem. Phys. 122, 084105 (2005).
- [77] C. P. Massen and J. P. K. Doye, Phys. Rev. E 71, 046101 (2005).
- [78] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. (MIT Press, Cambridge, MA, 2009).
- [79] D. J. Watts and S. H. Strogatz, Nature (London) 393, 440 (1998).
- [80] J. M. Carr and D. J. Wales, J. Phys. Chem. B 112, 8760 (2008).
- [81] A. Clauset, C. R. Shalizi, and M. E. Newman, SIAM Rev. 51, 661 (2009).
- [82] A. Cavalli, U. Haberthür, E. Paci, and A. Caflisch, Protein Sci. 12, 1801 (2003).
- [83] D. Gfeller, D. M. de Lachapelle, P. De Los Rios, G. Caldarelli, and F. Rao, Phys. Rev. E 76, 026113 (2007).
- [84] A. L. Barabási and R. Albert, Science 286, 509 (1999).
- [85] C. Beltrán, Constr. Approx. 37, 135 (2013).
- [86] D. Mehta, T. Chen, J. D. Hauenstein, and D. J. Wales, J. Chem. Phys. 141, 121104 (2014).
- [87] D. Mehta, T. Chen, J. W. Morgan, and D. J. Wales, J. Chem. Phys. 142, 194113 (2015).
- [88] T. Chen and D. Mehta, arXiv:1504.06622.
- [89] F. Rao and A. Caflisch, J. Mol. Biol. 342, 299 (2004).
- [90] C. Hughes, D. Mehta, and D. J. Wales, J. Chem. Phys. 140, 194104 (2014).