

Preferential attachment during the evolution of a potential energy landscape

Claire P. Massen

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Jonathan P. K. Doye

Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

(Received 21 June 2007; accepted 27 July 2007; published online 18 September 2007)

It has previously been shown that the network of connected minima on a potential energy landscape is scale-free, and that this reflects a power-law distribution for the areas of the basins of attraction surrounding the minima. Here, the aim is to understand more about the physical origins of these puzzling properties by examining how the potential energy landscape of a 13-atom cluster evolves with the range of the potential. In particular, on decreasing the range of the potential the number of stationary points increases and thus the landscape becomes rougher and the network gets larger. Thus, the evolution of the potential energy landscape can be followed from one with just a single minimum to a complex landscape with many minima and a scale-free pattern of connections. It is found that during this growth process, new edges in the network of connected minima preferentially attach to more highly connected minima, thus leading to the scale-free character. Furthermore, minima that appear when the range of the potential is shorter and the network is larger have smaller basins of attraction. As there are many of these smaller basins because the network grows exponentially, the observed growth process thus also gives rise to a power-law distribution for the hyperareas of the basins. © 2007 American Institute of Physics. [DOI: 10.1063/1.2773721]

I. INTRODUCTION

A potential energy landscape (PEL) is a high-dimensional surface describing how the potential energy of a system varies with the coordinates of each atom in the system.¹ Since PELs control the structure, thermodynamics, and dynamics of a system, an ongoing research program has sought to understand how the behavior of a system can be related back to features of a complex PEL. This approach has led to new insights into protein folding² and the nature of the glass transition.³ In these applications the emphasis is on how PELs differ, for example between proteins that are good and bad folders or liquids that are fragile and strong. In this paper, by contrast, we are interested in more universal properties of high-dimensional PELs.

When characterizing PELs the focus is often on the stationary points of these surfaces. Some fundamental properties associated with stationary points of PELs are well established. For example, the number of stationary points increases exponentially with system size.^{4–6} The energy distribution of both minima and transition states follows a Gaussian distribution, as has been seen empirically^{7,8} and predicted theoretically from the central limit theorem.⁹ More recently, network-derived insights have been obtained for PELs.^{10–14} To obtain a network representation of a PEL, each minima is mapped to a node in the network and each transition state to an edge, as in Fig. 1. The resulting “inherent structure” network^{10,11} provides a dynamically motivated description of the connectivity and organization of the PEL, since at sufficiently low temperatures a system vibrates

within the basin of attraction of a minimum, or inherent structure,¹⁵ before moving to another via a transition state valley.¹⁶

For small Lennard-Jones clusters, these networks have been shown to have both scale-free and small-world properties.^{10,11} The characteristic of a scale-free network is a power-law tail to the degree distribution,¹⁷ where the degree of a node is the number of edges connected to it, and the characteristic of a small-world network is a logarithmic increase in the average separation between nodes with network size.¹⁸ These two properties are related since the high-degree hubs present in scale-free networks connect up the network, making the path lengths between nodes very short.¹⁹ Recently, a wide variety of networks have been shown to be scale-free,^{20,21} including the worldwide web,^{22,23} social networks^{24,25} and biological networks.^{26,27} Typically, the exponent of the power law lies between -2 and -3 , as is the case for the inherent structure networks, which have an exponent of approximately -2.78 .^{10,11} Interestingly, both networks based on the configuration space of a protein,^{28,29} although defined in a more coarse-grained way than the inherent structure networks, and inherent structure networks for mean-field spin glasses³⁰ have also been found to be scale-free, suggesting that this feature could be a universal property of PELs.

A popular explanation for the scale-free character of many of these networks is in terms of preferential attachment during the growth of these networks.¹⁷ In preferential attachment, new nodes added to the network are more likely to attach to old nodes with high degree. Therefore, the degree

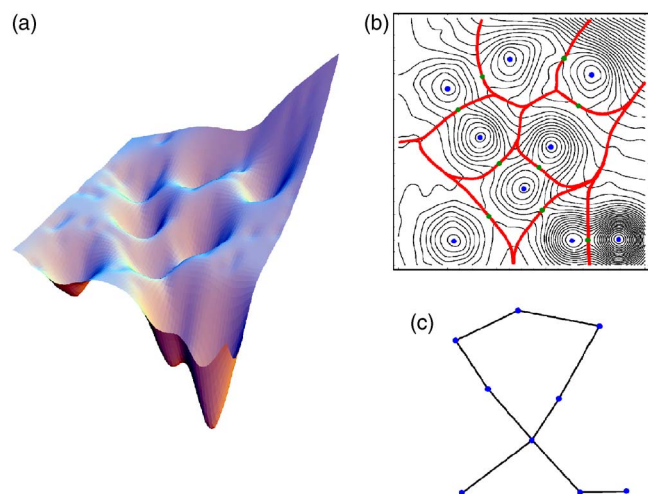


FIG. 1. (Color online) (a) A model two-dimensional energy landscape. (b) A contour plot of this surface illustrating the inherent structure partition of the configuration space into basins of attraction surrounding minima. The basin boundaries are represented by thick lines, and the minima and transition states by dots. (c) The resulting representation of the landscape as a network.

of high-degree nodes increases at a faster rate, a “rich get richer” growth process, and leads to a power-law degree distribution.

By contrast, a PEL is static—it is just determined by the form of the potential and the system size. Instead, for PELs, the scale-free nature seems to be related to the distribution of basin areas. Each minimum on the PEL is surrounded by a basin of attraction, a region of configuration space where following the gradient downhill will lead to the same minimum. Low-energy minima have large basins of attraction,^{31,32} and so unsurprisingly they have high degree because they can fit many neighbors along their long boundaries. A broad degree distribution thus reflects a broad basin area distribution.

This connection between area and degree is more explicit in Apollonian networks, which are model spatial, scale-free networks^{33,34} that have similar topological properties to inherent structure networks.^{13,34} These Apollonian networks are based on Apollonian packings in which all of space is tiled by hyperspheres (see Fig. 2 for a two-dimensional example). Each hypersphere gives rise to a node in the network, and two nodes are connected if the corresponding hyperspheres touch. The packing is constructed iteratively and at each step in the generation of the packing each interstice is filled by the largest possible hypersphere that just touches the surrounding hyperspheres. The fractal, self-similar character of the packings leads both to the scale-free topology of the Apollonian networks, and a power-law distribution for the hypervolumes of the hyperspheres.

The Apollonian packings and networks suggest how the basins of attraction on a PEL might have to tile configuration space in order to generate a scale-free inherent structure network. Namely, the tiling would need to be hierarchical with larger basins surrounded by smaller basins, which are in turn surrounded by smaller basins, and so on. Interestingly, when the basin area distribution has been probed for model liquids, a power-law distribution has been found with the exponent

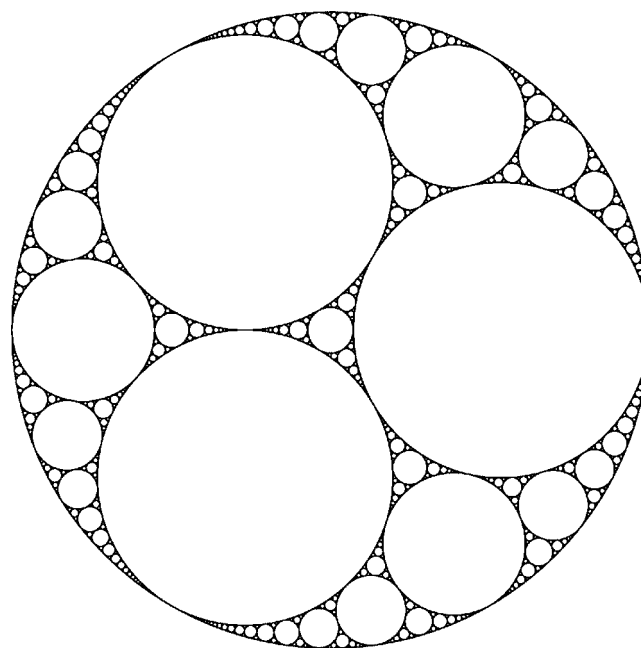


FIG. 2. A two-dimensional Apollonian packing. Space is filled with different-sized disks, starting from an initial configuration where the three larger disks are placed within the bounding circle. The packing is generated iteratively, with the largest disk possible added to each gap at each iteration. This process is continued *ad infinitum*, thus filling the space with successively smaller and smaller disks.

expected from analogy to the Apollonian packings,¹² thus suggesting that the inherent structure networks for these systems are scale-free (note that it is effectively impossible to obtain these networks directly due to the high dimensionality of their configuration space). Although these results and ideas show how the topology of the inherent structure networks can be understood in terms of how the basins of attraction tile configuration space, why this tiling should be hierarchical, fractal, and Apollonian-like is still a puzzle.

Recently, using a simple model we explored whether these features could simply be driven by the variation in the energies of the minima.³⁵ In this “eggbox” model, we took the Gaussian distribution of energies of the minima as a starting point, and examined how the PEL landscape evolved (e.g., through the swallowing up of higher energy minima by deep basins) as the variance of the Gaussian increased. Although the resulting area distribution was considerably broader than for a flat landscape where all the minima have the same energy, these changes were not sufficient to generate a power-law tail for the degree or basin area distributions.

Here, we provide an alternative perspective on why the inherent structure networks for PELs are scale-free, and, as we will see, hence obtain insights into the nature of the division of configuration space into basins of attraction. Although, as already mentioned, a PEL is static, there may be ways to envisage the generation of the network as a quasi-growth process, and hence to apply the ideas of preferential attachment. Recently, we looked at the role played by the energetic ordering of minima.¹⁴ Here, we wish to look at how the PEL (and the inherent structure network) evolves on going from a simple landscape with one minimum to a complex one with many (N) minima and a scale-free topology.

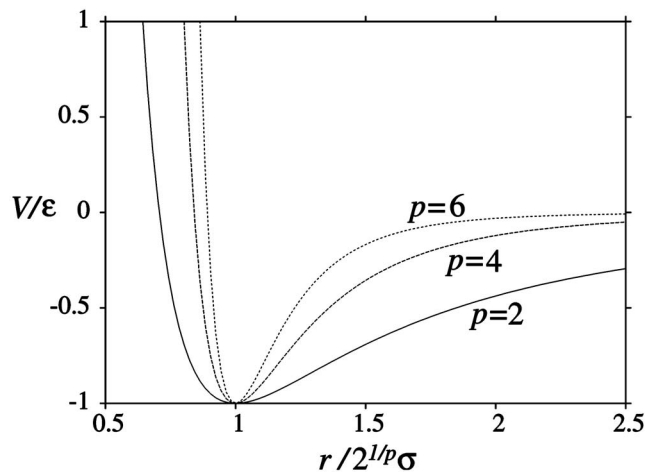


FIG. 3. The generalized Lennard-Jones potential [Eq. (1)] for $p=2, 4$, and 6 .

The reverse of this process, smoothing a complex landscape so that there are significantly fewer minima, and perhaps only one, has been a common approach in global optimization.^{36–44} The idea behind this approach is that global optimization is likely to be much easier on the smoothed landscape, and trivially so if there is only one minimum left. Of this set of global optimization algorithms, particularly relevant to us are those that achieve this smoothing by changing a potential parameter.^{37–39,41–44}

Here, we examine the evolution of the landscape of the 13-atom cluster for a generalized Lennard-Jones potential with parameter p , defined by

$$V = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{2p} - \left(\frac{\sigma}{r} \right)^p \right], \quad (1)$$

where r is the distance between two atoms, ϵ is the well-depth and $2^{1/p}\sigma$ is the equilibrium pair separation. This is summed over all pairs of atoms to give the total potential energy of a configuration. The range of the potential is determined by the parameter p , where for $p=6$ the potential is equivalent to the Lennard-Jones potential. As p decreases, the repulsion becomes less steep and the attraction has a longer range. The net effect is that the width of the well increases, as shown in Fig. 3.

The effects of p (or the analogous parameter for the Morse potential) on the landscape have been well studied.^{37,45–53} At sufficiently small p , there is a single minimum on the landscape. As p increases, the number of minima increases rapidly, because potential wells surrounding the minima become narrower, allowing new minima to appear at higher energy. Equivalently, as p decreases, the lower energy wells become wider, thus swallowing up the higher energy minima. Downhill barriers also increase as p increases. The net effect is that a smooth landscape at low p becomes rougher.

In Sec. II, we provide details of how the appearance and disappearance of minima and transition states can be described by catastrophe theory,^{54,55} and the effect on the inherent structure network. In Sec. III, we describe the method used to detect preferential attachment. In Sec. IV, we discuss

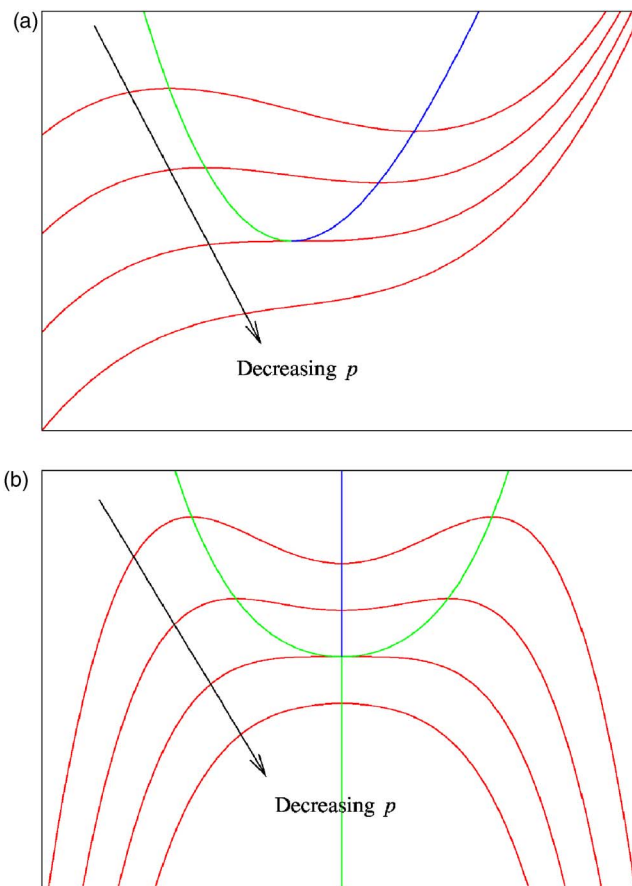


FIG. 4. (Color online) Changes in a landscape close to a catastrophe. The loci of the stationary points are also plotted. (a) The fold catastrophe. Initially, there are two stationary points, a maximum and a minimum. Upon varying a parameter, decreasing p (increasing the range) in the case of the PEL, these stationary points move closer together until they collide and disappear. (b) The cusp catastrophe. Initially, there is one minimum and two maxima, which collide upon increasing the range to leave a single maximum.

how the landscape for the 13-atom Lennard-Jones cluster evolves, and whether preferential attachment provides a useful description of that evolution.

II. CATASTROPHES

The two most common types of catastrophe for PELs are the fold and the cusp, illustrated in Fig. 4. In the fold, a minimum and a maximum move closer together as some parameter is varied, in this case as the range is increased. At the catastrophe the two stationary points meet and disappear, leaving just a shoulder in the landscape. The cusp in Fig. 4(b) involves two symmetry-related maxima and one minimum. Upon increasing the range, they move closer together until they collide, leaving one maximum.

Figure 4 is a one-dimensional picture, whereas the PEL is high-dimensional. However, these catastrophes can provide a good description of the evolution of one-dimensional “reaction coordinates” that link stationary points on the PEL.⁵⁴ These reaction coordinates are defined in terms of steepest-descent pathways, i.e., the paths follow the gradient downhill at every point, and thus there is no component of the gradient in the space orthogonal to the coordinate. Hence,

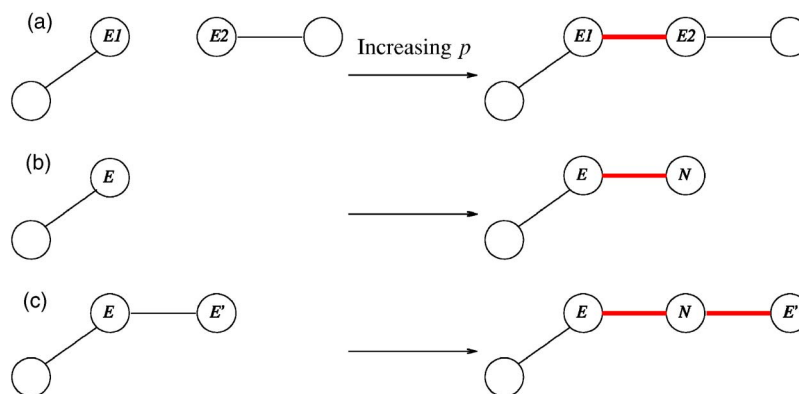


FIG. 5. (Color online) Changes to the inherent structure network that can occur on increasing p . (a) Addition of an internal edge (shown in bold) connecting two existing nodes, $E1$ and $E2$. (b) Addition of an external edge connecting a new node N to an existing node E . (c) Addition of an external edge via a cusp. The initial transition state connects two permutational isomers, E and E' , forming a self-connection. The new node is inserted into this edge, forming connections to both permutational isomers. However, because we ignore ME/SCs, the overall effect on the network topology is equivalent to that in (b).

the points where there is a maximum or minimum in the reaction coordinate correspond to stationary points of the PEL.

We are interested in catastrophes involving minima and transition states, where minima are stationary points that are minimal in all directions, and transition states are maximal in one direction and minimal in all others, i.e., the Hessian has one negative eigenvalue. This leaves two possible scenarios. In the first, the minima in Fig. 4 represent minima and the maxima transition states. In the second, the minima in Fig. 4 represent transition states and the maxima index 2 saddles, which are maximal in one more dimension than transition states. The cusp catastrophe can only occur when there is symmetry present, as the pathways on either side of the central stationary point are equivalent. Therefore, any stationary points connected by the pathways must be permutational isomers, where the system has the same structure but some atoms swap positions. The low- p maximum must therefore connect two permutational isomers of the same minimum.

There are some subtleties in the way the inherent structure network has been defined. Edges correspond to transition states, but not all transition states give rise to edges. More than one transition state may connect the same pair of nodes, forming a multiple edge (ME), and some connect permutational isomers of the same minimum, forming a self-connection (SC). ME/SCs are not considered in the inherent structure network, as we are interested only in whether two minima are connected, not how many times, and we do not distinguish between permutational isomers of a minimum. This choice is important for the network topology.^{11,56}

The two catastrophes shown in Fig. 4 can affect the inherent structure network as shown in Fig. 5. We classify catastrophes as the addition of either an internal or external edge to the network.^{57–60} An internal edge involves the addition of an edge between two existing nodes. This can occur on increasing p via a fold catastrophe, creating a transition state and an index 2 saddle from an inflection point, or a cusp catastrophe, creating one transition state and two index 2 saddles from an index 2 saddle. Both these processes occur along the boundaries between two basins of attraction. The new transition state may connect two permutational isomers

of the same minimum, in which case it would form a SC, or it may connect two minima that are already connected by a transition state, forming a ME, and therefore have no effect on the network topology.

An external edge involves the addition of a new node and one new edge. This can either be from a fold catastrophe, an inflection point becoming a minimum and a transition state, or a cusp catastrophe, a transition state becoming a minimum and two transition states. Symmetry restrictions on the cusp catastrophe affect the network as shown in Fig. 5(c). The initial transition state connects two permutational isomers; therefore, it forms a SC in the network and is not included in our representation. A cusp inserts a minimum into this edge that is connected to both permutational isomers. As we only count one permutational isomer in the inherent structure network, the effect on the network topology is that one external edge is added.

When a new minimum appears on the landscape, an external edge is formed to one existing node by the new transition state that was involved in the catastrophe. However, the change to the network topology can be more complicated. The basin of attraction associated with a new minimum generated by a fold catastrophe has a finite area, and as such can cause reaction pathways from existing transition states to be rewired to connect to it if the existing reaction pathway passed through the region of configuration space now assigned to the basin of the new minimum (Fig. 6). The new node can therefore have initial degree greater than 1, although there is only one new transition state. These rewired transition states may be rewired from MEs and as such appear to be new edges in terms of the network topology. Because such new edges connect to the new node, they are classified as external edges.

Although these are the most likely types of catastrophe, there may be other, more complicated catastrophes. However, such catastrophes involve higher symmetries and are therefore rare. Furthermore, we do not search specifically for folds and cusps, simply for stationary points on the PEL. Starting from $p=6$, for which it is likely that all of the stationary points are known, we decrease p (increasing the range of the potential) and record when stationary points

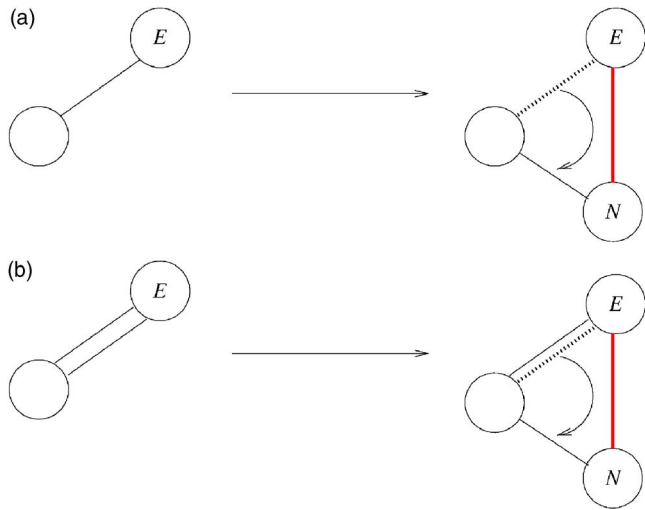


FIG. 6. (Color online) The two major ways that the addition of an external edge (shown in bold) connecting new node N to existing node E can lead to an edge being rewired, thus increasing the initial degree of the new node. In (a), the edge is rewired from E , thus reducing its degree. In (b), the rewired edge was originally a ME; thus, the degree of E is not affected.

disappear from the surface, and when transition states rewire to connect to different minima. This information is then inverted to give the details of the network's growth. While it is possible for a minimum or transition state to appear as the range is increased, this type of event is difficult to detect and is expected to be relatively rare, so it is not included. Rewiring of transition states is only considered when a minimum is involved in a catastrophe, as this is the most likely occasion for it to occur.

III. MEASURING PREFERENTIAL ATTACHMENT

In preferential attachment models, when a new node appears in the network, it is assumed to connect to an existing node with probability Π , depending on its degree k as $\Pi \propto k^\alpha$. The exponent α describes the strength of the preferential attachment. In the original Barabási-Albert model,¹⁷ α must be 1 to give a scale-free network.⁶¹ In more complex models, allowing for example new edges between a pair of existing nodes,^{62–67} other values of α can give rise to scale-free networks.

The exponent α can potentially be determined from time-resolved data for the growth of the network. However, extracting the exponent is not straightforward for two reasons. First, network growth is typically stochastic. A given preferential attachment rule can result in many different networks, and as such the exponent cannot be determined exactly. Second, at each time step the probability of connecting to a node with degree k is $\Pi = k^\alpha / \sum_i k_i^\alpha$. The normalizing sum in the denominator k is time dependent, and depends on the unknown exponent. Approximations have been used previously, such as comparing the network at two different stages of its growth and assuming the normalizing sum remains constant over this time period,^{63,64,66} or assuming that the normalization sum is proportional to the size of the network.⁶² By contrast, we use an iterative, self-consistent

method that we recently introduced.¹⁴ In this method, we estimate α in order to evaluate the sum, then use the results to obtain a better estimate for α .

The number of edges gained by all nodes with degree k at time t , $\Delta k(k, t)$, depends on the number of nodes with degree k at time t , $n_k(k, t)$, and is given by

$$\Delta k(k, t) = n_k(k, t) \Pi(k, t) = \frac{n_k(k, t) k^\alpha}{\sum_i k_i(t)^\alpha} = \frac{n_k(k, t) f(k)}{c_t(t)}, \quad (2)$$

where we define $f(k) = k^\alpha$ and $c_t(t) = \sum_i k_i(t)^\alpha$. $f(k)$ is obtained by summing both sides over all time steps, thus avoiding problems from steps where $n_k(k, t) = 0$, giving

$$f(k) = \frac{\sum_t \Delta k(k, t)}{\sum_t n_k(k, t) / c_t(t)}. \quad (3)$$

α is then obtained from a linear least-squares fit to $\log f(k) = \alpha \log k$, weighted by the number of terms in the sum in the denominator of Eq. (3). Logarithmic binning over k reduces errors from degrees where $\sum_t \Delta k(k, t) = 0$, which cannot be included in a log-log plot. It has been shown that the resulting exponent is independent of the initial estimate for α , and the method generally converges.¹⁴ In tests on model networks grown with a known preferential attachment rule, the exponent obtained had an absolute error less than approximately 0.13 for $0.5 \leq \alpha \leq 1.5$, with a much lower error for α closer to 1.

We investigate the addition of internal and external edges independently. For internal edges, the probability of a pair of nodes gaining an edge is assumed to depend on the product of their degrees as $\Pi(k_i k_j) \propto (k_i k_j)^\alpha$. If multiple edges are not allowed in a network, when determining α only those pairs of nodes that are not connected should be used in the normalization sum $c_t(t)$. This modification has been seen to greatly improve the exponent obtained, reducing the error from approximately 20% to 1% for the $\alpha = 1$ case.¹⁴ This approach can straightforwardly be applied to the evolution of

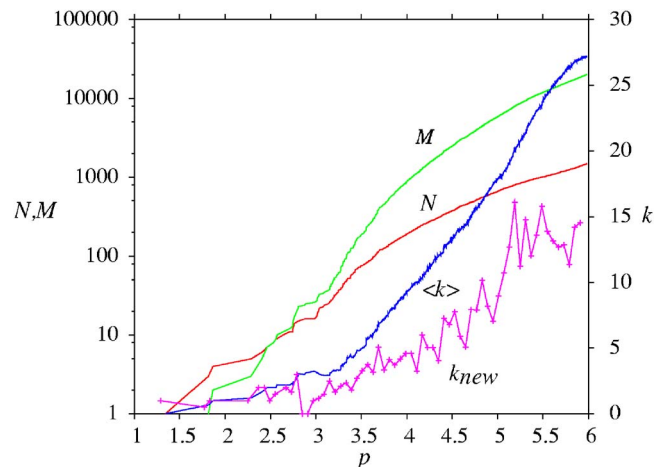


FIG. 7. (Color online) Growth with p of the number of nodes, N , the number of edges, M , the average degree, $\langle k \rangle = 2M/N$, and the initial degree of a new node, k_{new} (binned over p).

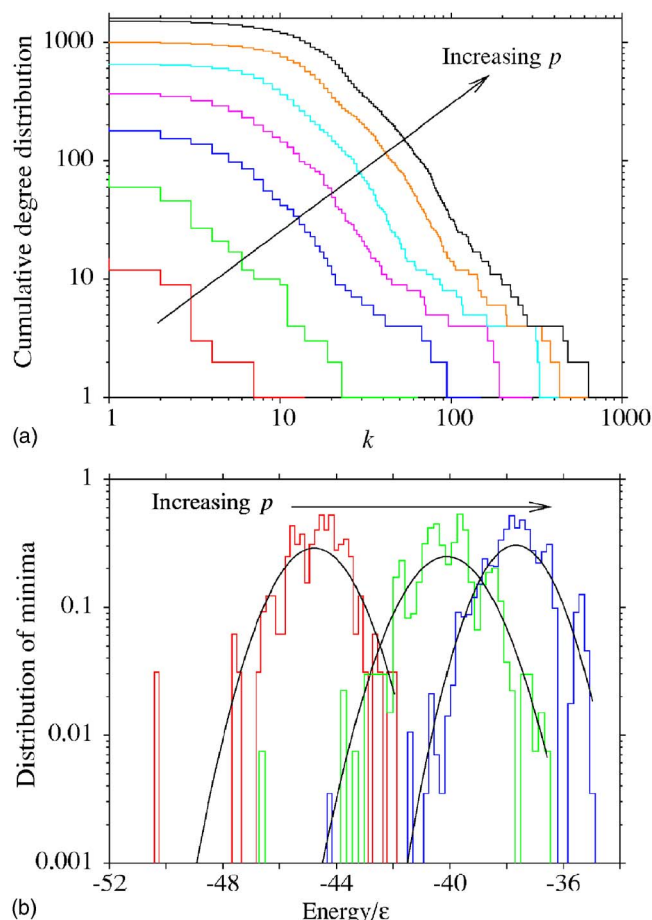


FIG. 8. (Color online) (a) Cumulative degree distribution (i.e., the number of nodes with degree greater than k) and (b) distribution of the energies of the minima at various values of p during the growth of the network. In (b), Gaussian best fits are also shown. In calculating these best fits, the global minimum is excluded as it has especially low energy, due to the particularly stable icosahedral structure of the cluster.

the inherent structure networks, the only unusual feature being that the networks grow as a function of the potential parameter p , not time.

IV. RESULTS

A. Network properties

The number of nodes and edges increase as the range parameter is increased toward $p=6$, at which point the network contains 1509 nodes and 20 691 edges. As shown in Fig. 7, the number of nodes increases roughly exponentially with p , i.e., the network grows faster as p approaches 6. The number of edges also increases roughly exponentially with p . As cusp catastrophes occur at either transition states or index 2 saddles, and fold catastrophes occur at inflection points, both of these will be more likely when the landscape is rougher (i.e., the network is bigger). In particular, if $dN/dp \propto N$ exponential growth results.

The number of edges increases faster than the number of nodes, so the average degree increases, i.e., the network is accelerating and becoming better connected. This is due to two factors: first, the addition of internal edges. Second, the initial degree of a new node increases with p . If a node is

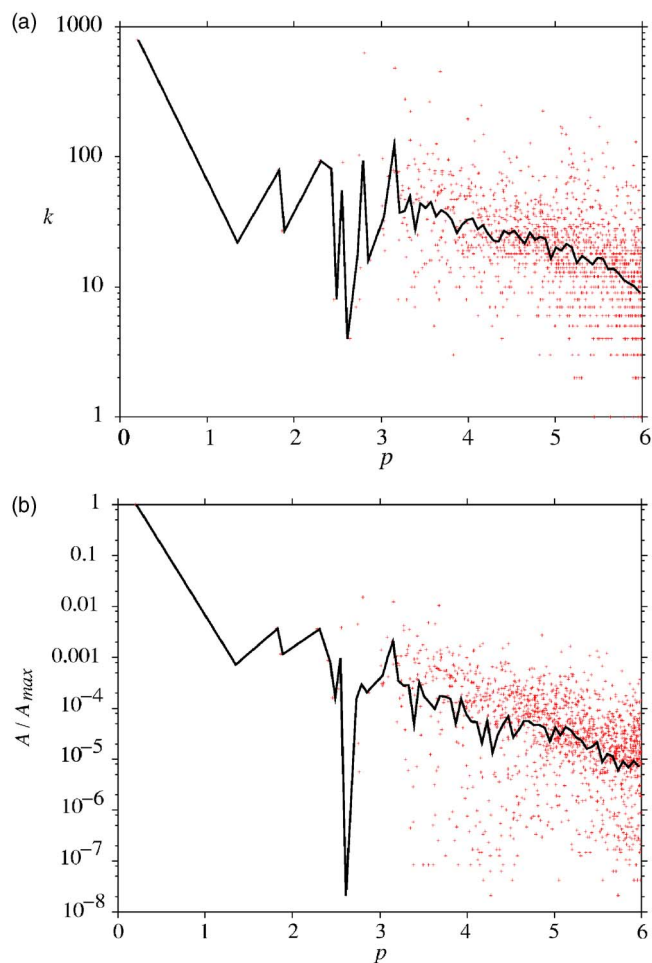


FIG. 9. (Color online) Properties of a node at $p=6$ as a function of the value of p when it was created. (a) Node degree, k , and (b) basin area, A . Individual points correspond to individual nodes, and the solid line is a binned geometric average. Note that, due to the exponential growth of the network, only 17 nodes, 1.1% of the total, appear at $p \leq 3$, where the correlation of the properties with p is weaker.

added at higher p , it is likely to link to more existing nodes because there are more nodes in the network, and more edges to rewired. Some of these edges are rewired from MEs, thus increasing the average degree.

The degree distribution has a similar form throughout the growth process, as shown in Fig. 8(a), indicating that the growth is fairly homogeneous in terms of network topology, and the network reaches a kind of steady state. Furthermore, the distribution of energies of the minima is approximately Gaussian throughout the growth process [Fig. 8(b)]. Although the initial degree of “newer” nodes (those added at higher p) is higher, the final degree at $p=6$ is lower than for the older nodes [Fig. 9(a)]. This “age” dependence of the degree is indicative of preferential attachment and has been seen previously, for example in a protein interaction network.⁶⁵ It should be noted that the network is only large enough to study degree distributions for $p \geq 3$. In this region, the growth of the number of stationary points, and particularly the average degree, is also fairly steady (Fig. 7).

B. Preferential attachment

The exponents that we obtained by applying the algorithm in Sec. III are shown in Table I. They are close to 1,

indicating that preferential attachment is present in the growth of the network. New edges are more likely to link to existing nodes with high degree, leading to a scale-free network. The exponent is slightly greater than 1 for external edges, both for those associated with new transition states and those associated with rewired MEs. When a new transition state appears on the PEL that gives rise to an external edge, the degree of the existing node to which it connects increases by 1. Simultaneously, some edges are rewired from existing nodes, increasing the initial degree of the new node. However, due to the spatial nature of the PEL, these rewired edges are likely to be rewired from the existing node that is connected to the new transition state (see Fig. 6). In fact, on average, when such an external edge is added, the existing node at the end of the edge loses six edges to rewiring, while only gaining one. This opposes preferential attachment by taking edges from high-degree nodes. However, this effect is canceled out by the preferential attachment associated with external edges that simultaneously appear due to the rewiring of MEs. Furthermore, the frequency with which external edges are added is lower than that for internal edges (Table I) and the appearance of new minima increases the available basin boundary along which new internal edges can be generated.

Internal edges do increase the degree of high-degree nodes, as expected for preferential attachment, and the exponent for this process is very close to 1. Since these “time steps” dominate the network growth, the preferential attachment associated with the addition of internal edges seems to be the main driving force for the scale-free topology of the final network. As internal edges are generated by catastrophes occurring along the basin boundaries, the larger, older minima are likely to garner more internal edges, hence the preferential attachment. Indeed, Fig. 10(a) shows that the degree and basin area are strongly correlated.

To probe the connection to the basin area distribution in more detail, we have measured the areas of the basins of attraction directly. This task is achieved simply by choosing random points in configuration space and then performing a local minimization. The area of a basin is then proportional to the number of points that lead to the corresponding minimum after minimization. Newer minima have smaller basins of attraction in the final PEL at $p=6$ [Fig. 9(b)]. When new minima appear on the landscape they are likely to do so near the edge of a basin of attraction. Near the center of the basin, the basin behaves harmonically, as quadratic terms dominate

TABLE I. Exponents found for preferential attachment of external edges ($\Pi(k) \propto k^\alpha$) and internal edges ($\Pi(k_i k_j) \propto (k_i k_j)^\alpha$). The number of data points used to obtain the exponents is also shown. External edges are further broken down into those that involve addition of a new transition state and those that involve rewiring a transition state. Rewired transition states are only considered if they involve an increase in degree of the existing node, i.e., before the catastrophe, the transition state formed a ME.

Type of time step	Exponent α	Number
External edges (new TS)	1.20	1 199
External edges (rewired TS)	1.24	7 275
Internal edges	0.99	12 274

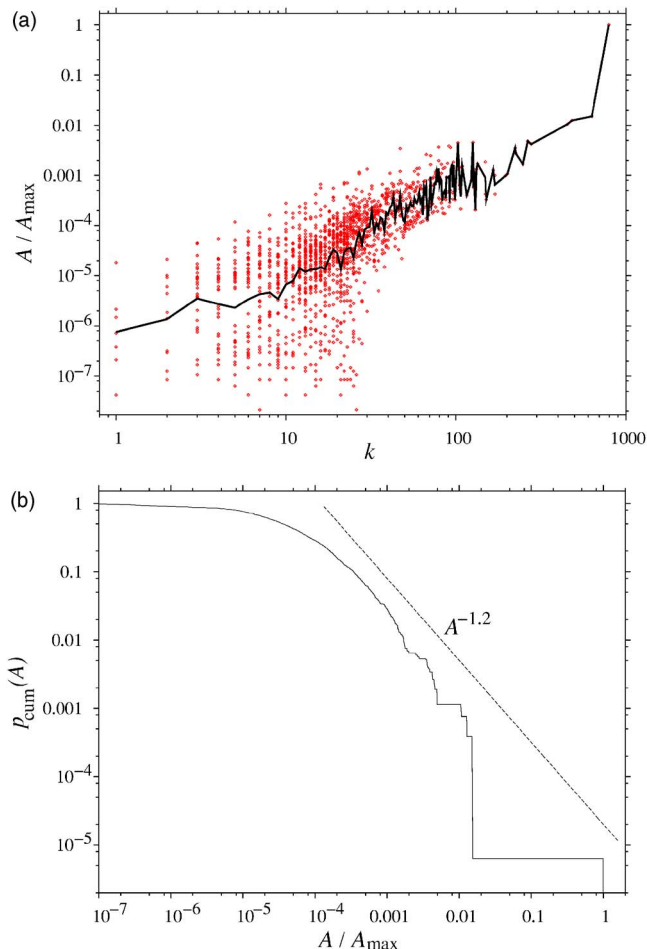


FIG. 10. (Color online) (a) Degree dependence of the area of the basin of attraction of a minimum for LJ₁₃. Individual points correspond to individual minima, and the solid line is a geometric average. (b) The cumulative basin area distribution. In (b) a power-law form corresponding to $p(A) \sim A^{-2.2}$ (the exponent of the cumulative distribution is greater by 1) has also been plotted.

the Taylor expansion of the energy about the minimum. Catastrophes are driven by higher-order anharmonic terms, and so will only occur at sufficient distance from the minimum.

At the exact moment at which the catastrophe occurs and the minimum appears in the network, the minimum and the new transition state also involved in the catastrophe are at the same point. For the fold catastrophe, the basin associated with this minimum initially has zero extent in the direction of the transition state, but both in the opposite direction and all those directions perpendicular to this reaction coordinate, it is of finite extent. As p increases, the distance between the transition state and minimum will increase, albeit relatively slowly—close to the catastrophe this distance can be approximated by $\sqrt{6\Delta V/\lambda}$, where ΔV is the barrier, and λ is the curvature at the minimum and transition state along the reaction path.⁵⁴ Therefore, the area of the new basin will probably initially increase, but it is always likely to be smaller than the basins of the older minima between which it is sandwiched. Hence, as p increases and there are more minima on the landscape, the new minima will have increasingly small basins, as they are more constrained by the more closely spaced basins.

Combining this effect with the growth of the network leads to a power-law area distribution. New basins that appear at higher p have smaller basin areas. Due to the exponential increase of N with p , there are many of these new, small basins. Furthermore, the nature of the catastrophes means that the new basins are likely to appear at or near the boundaries of larger, existing basins. For this reason, the distribution of minima in configuration space will be inhomogeneous, with minima concentrated around the edges of the larger basins in a manner that is somewhat similar to the spatial distribution of hyperspheres in the Apollonian packing (Fig. 2).

Both the number of minima at a given p and the area of a basin (at $p=6$) that appeared at p show an approximately exponential dependence on p , i.e., $N(p) \sim \exp(\nu p)$ and $A(p) \sim \exp(-\alpha p)$. Hence, we can derive an expression for the area distribution. The number of basins with area A at $p=6$ is given by

$$n(A)dA = \frac{dN}{dp} dp, \\ n(A) \sim \nu \exp(\nu p) \left| \frac{dp}{dA} \right| \sim \frac{\nu \exp(\nu p)}{\alpha \exp(-\alpha p)}. \quad (4)$$

Substituting for p via $A(p) \sim \exp(-\alpha p)$, we arrive at

$$n(A) \sim \frac{\nu}{\alpha} A^{-1-\nu/\alpha}. \quad (5)$$

This mechanism for generating a power law has been termed “combination of exponentials” by Newman.⁶⁸ The observed exponents are $\nu \approx 1.1$, $\alpha \approx 1.2$, giving $P(A) \sim A^{-1.9}$. This result is in line with the observed approximate power-law tail to the basin area distribution for LJ₁₃ [Fig. 10(b)]. The exponent is also close to that for the area distributions observed for liquids and for Apollonian packings.¹² For the latter the exponent is exactly -2 in the limit of high dimensionality. Interestingly, the analogies to the Apollonian packing go farther, as the power law in the Apollonian case can also be considered to arise from a combination of exponentials. Namely, for the two-dimensional Apollonian packing, it has been shown that both the number of disks and their area depend exponentially on the number of iteration steps used to generate the packing, and that, furthermore, new disks preferentially attach to those with high degree.³⁴

V. CONCLUSION

We have used techniques from network theory to study the evolution of a PEL on going from a simple form with a single minimum to a complex landscape with a large number of minima as the range of the underlying interatomic potential is varied. In particular, we have examined how the inherent structure network, which provides a picture of the connectivity of the PEL, changes. New edges in this network are seen to preferentially attach to minima with a larger number of connections, helping to explain the scale-free nature of the inherent structure network. This preferential attachment is very close to linear, and can be understood by considering the underlying catastrophes that lead to the appearance of

new minima and transition states on the PEL. Each catastrophe that occurs, as well as increasing the degree of a minimum, also makes the surface rougher, leading to further catastrophes. These catastrophes are also more likely to occur at or near boundaries; therefore, they are more likely to occur near large basins, which have high degree.

Consideration of this growth process leads to further insight into the nature of the division of the PEL into basins of attraction. Previous work has shown that, similar to the Apollonian packing, PELs have power-law distributions for the hyperareas of these basins of attraction, suggesting that the basins tile the landscape in a fractal-like manner. The current results are able to shed light on this puzzling character of the PEL. During the evolution of the PEL small new basins appear near the boundaries of larger existing boundaries, as in the generation of the Apollonian packing. The network grows exponentially, because catastrophes make the network larger and the surface rougher, enabling further catastrophes. The combination of this exponential growth and the smaller size of newer basins leads to a power-law area distribution, in agreement with observed results for PELs.

To continuously transform the PEL to one with a single minimum, we chose to vary the exponents in the Lennard-Jones potential. However, this is not the only way that one could envisage achieving this aim. Indeed, a variety of approaches to transform the PELs of Lennard-Jones clusters has been explored in the global optimization literature.^{37,39,41–44} However, we would expect the way that the PEL and inherent structure network evolves under these different transformation to be similar to that which we have observed here, since the key features of that evolution will apply in all cases. Namely, the appearance of stationary points is described by catastrophe theory, and these catastrophes preferentially occur near or on the boundaries of the basins of attraction.

ACKNOWLEDGMENTS

The authors would like to thank the Royal Society (J.P.K.D.) and the EPSRC(UK) for financial support (C.P.M.).

¹D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, 2004).

²J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, *Proteins* **21**, 167 (1995).

³P. G. Debenedetti and F. H. Stillinger, *Nature* **410**, 259 (2001).

⁴C. J. Tsai and K. D. Jordan, *J. Phys. Chem.* **97**, 11227 (1993).

⁵J. P. K. Doye and D. J. Wales, *J. Chem. Phys.* **116**, 3777 (2002).

⁶F. H. Stillinger, *Phys. Rev. E* **59**, 48 (1999).

⁷S. Büchner and A. Heuer, *Phys. Rev. E* **60**, 6507 (1999).

⁸F. Sciortino, W. Kob, and P. Tartaglia, *Phys. Rev. Lett.* **83**, 3214 (1999).

⁹A. Heuer and S. Büchner, *J. Phys.: Condens. Matter* **12**, 6535 (2000).

¹⁰J. P. K. Doye, *Phys. Rev. Lett.* **88**, 238701 (2002).

¹¹J. P. K. Doye and C. P. Massen, *J. Chem. Phys.* **122**, 084105 (2005).

¹²C. P. Massen and J. P. K. Doye, *Phys. Rev. E* **71**, 046101 (2005).

¹³J. P. K. Doye and C. P. Massen, in *Complexity, Metastability and Nonextensivity*, edited by C. Beck, G. Benedek, A. Rapisarda, and C. Tsallis (World Scientific, Singapore, 2005), pp. 375–384.

¹⁴C. P. Massen and J. P. K. Doye, *Physica A* **377**, 351 (2007).

¹⁵F. H. Stillinger and T. A. Weber, *Science* **225**, 983 (1984).

¹⁶T. B. Schröder, S. Sastry, J. C. Dyre, and S. C. Glotzer, *J. Chem. Phys.* **112**, 9834 (2000).

¹⁷A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).

- ¹⁸D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
- ¹⁹R. Cohen and S. Havlin, *Phys. Rev. Lett.* **90**, 058701 (2003).
- ²⁰M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).
- ²¹S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
- ²²R. Albert, H. Jeong, and A.-L. Barabási, *Nature* **401**, 130 (1999).
- ²³M. Faloutsos, P. Faloutsos, and C. Faloutsos, *Comput. Commun. Rev.* **29**, 251 (1999).
- ²⁴M. E. J. Newman, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 404 (2001).
- ²⁵S. Redner, *Eur. Phys. J. B* **4**, 131 (1998).
- ²⁶H. Jeong, B. Tomber, R. Albert, Z. N. Oltvai, and A.-L. Barabási, *Nature* **407**, 651 (2000).
- ²⁷H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, *Nature* **411**, 41 (2001).
- ²⁸F. Rao and A. Caffisch, *J. Mol. Biol.* **342**, 299 (2004).
- ²⁹D. Gfeller, P. De Los Rios, A. Caffisch, and F. Rao, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1817 (2007).
- ³⁰Z. Burda, A. Krzywicki, and O. C. Martin, arXiv:0707.1965.
- ³¹J. P. K. Doye, D. J. Wales, and M. A. Miller, *J. Chem. Phys.* **109**, 8143 (1998).
- ³²C. P. Massen and J. P. K. Doye, *Phys. Rev. E* **75**, 037101 (2007).
- ³³J. S. Andrade, Jr., H. J. Herrmann, R. F. S. Andrade, and L. R. da Silva, *Phys. Rev. Lett.* **94**, 018702 (2005).
- ³⁴J. P. K. Doye and C. P. Massen, *Phys. Rev. E* **71**, 016128 (2005).
- ³⁵C. P. Massen and J. P. K. Doye, *Physica A* **382**, 683 (2007).
- ³⁶F. H. Stillinger and T. A. Weber, *J. Stat. Phys.* **52**, 1429 (1988).
- ³⁷F. H. Stillinger and D. K. Stillinger, *J. Chem. Phys.* **93**, 6106 (1990).
- ³⁸T. Head-Gordon, F. H. Stillinger, and J. Arrecis, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11076 (1991).
- ³⁹J. Pillardy, K. A. Olszewski, and L. Piela, *J. Phys. Chem.* **96**, 4337 (1992).
- ⁴⁰L. Piela, K. A. Olszewski, and J. Pillardy, *J. Mol. Struct.: THEOCHEM* **308**, 229 (1994).
- ⁴¹J. Pillardy and L. Piela, *J. Phys. Chem.* **99**, 11805 (1995).
- ⁴²C.-S. Shao, R. H. Byrd, E. Eskow, and R. B. Schnabel, *J. Glob. Optim.* **16**, 167 (2000).
- ⁴³M. Locatelli and F. Schoen, *Comput. Optim. Appl.* **21**, 55 (2002).
- ⁴⁴J. P. K. Doye, *Phys. Rev. E* **62**, 8753 (2000).
- ⁴⁵M. R. Hoare and J. McInnes, *Faraday Discuss.* **61**, 12 (1976).
- ⁴⁶M. R. Hoare and J. A. McInnes, *Adv. Phys.* **32**, 791 (1983).
- ⁴⁷P. A. Braier, R. S. Berry, and D. J. Wales, *J. Chem. Phys.* **93**, 8745 (1990).
- ⁴⁸J. P. K. Doye, D. J. Wales, and R. S. Berry, *J. Chem. Phys.* **103**, 4234 (1995).
- ⁴⁹J. P. K. Doye and D. J. Wales, *Science* **271**, 484 (1996).
- ⁵⁰J. P. K. Doye and D. J. Wales, *J. Phys. B* **29**, 4859 (1996).
- ⁵¹J. P. K. Doye and D. J. Wales, *J. Chem. Soc., Faraday Trans.* **93**, 4233 (1997).
- ⁵²M. A. Miller, J. P. K. Doye, and D. J. Wales, *Phys. Rev. E* **60**, 3701 (1999).
- ⁵³M. A. Miller, J. P. K. Doye, and D. J. Wales, *J. Chem. Phys.* **110**, 328 (1999).
- ⁵⁴D. J. Wales, *Science* **293**, 2067 (2001).
- ⁵⁵R. Gilmore, *Catastrophe Theory for Scientists and Engineers* (Wiley, New York, 1981).
- ⁵⁶S. Maslov, K. Sneppen, and A. Zaliznyak, *Physica A* **333**, 529 (2004).
- ⁵⁷R. Albert and A.-L. Barabási, *Phys. Rev. Lett.* **85**, 5234 (2000).
- ⁵⁸S. N. Dorogovtsev and J. F. F. Mendes, *Phys. Rev. E* **63**, 025101(R) (2001).
- ⁵⁹J. S. Mattick and M. J. Gagen, *Science* **307**, 856 (2005).
- ⁶⁰M. J. Gagen and J. S. Mattick, *Phys. Rev. E* **72**, 016123 (2005).
- ⁶¹P. L. Krapivsky, S. Redner, and F. Leyvraz, *Phys. Rev. Lett.* **85**, 4629 (2000).
- ⁶²M. E. J. Newman, *Phys. Rev. E* **64**, 025102(R) (2001).
- ⁶³A.-L. Barabási, H. Jeong, Z. Nédá, E. Ravasz, A. Schubert, and T. Vicsek, *Physica A* **311**, 590 (2002).
- ⁶⁴H. Jeong, Z. Nédá, and A.-L. Barabási, *Europhys. Lett.* **61**, 567 (2003).
- ⁶⁵E. Eisenberg and E. Y. Levanon, *Phys. Rev. Lett.* **91**, 138701 (2003).
- ⁶⁶C. Roth, in ISWC-2005 Workshop on Semantic Network Analysis (2005), p. 29.
- ⁶⁷M. Peltomäki and M. Alava, *J. Stat. Mech.: Theor. Exp.* **2006**, P01010 (2006).
- ⁶⁸M. E. J. Newman, *Contemp. Phys.* **46**, 323 (2005).