Connectivity graph: Multiple connectivity on potential energy surface does affect the dynamics

T. Okushima,* T. Niiyama, K. S. Ikeda, and Y. Shimizu
Department of Physics, Ritsumeikan University, Noji-higashi 1-1-1, Kusatsu 525-8577, Japan
(Received 7 September 2006; revised manuscript received 30 July 2007; published 13 September 2007)

A mapping called a *connectivity graph* is proposed to visualize the complicated networks between local minima via saddles on many-dimensional potential energy surfaces. With this mapping, the effect of the topography of a complex potential energy surface on the dynamics is studied in model funnel potentials and a Lennard-Jones cluster. We present strong evidence that energetically expensive but dynamically relevant saddles are indispensable for kinetic dynamics. The results show that the connectivity graphs provide a sound basis for understanding nonequilibrium dynamics.

DOI: 10.1103/PhysRevE.76.036109 PACS number(s): 82.20.Kh, 31.50.−x, 89.75.Hc

Structural, thermodynamic, and dynamical properties observed in simulations and experiments on complicated systems, such as atomic microclusters [1–3], carbon nanostructures [4], biomolecules [5,6], and glass-forming materials [7–12], are attributed to the complicated geometry of the underlying potential energy surfaces (PESs) [13,14].

Because a PES is a manifold spanned in a many-dimensional space, it is hard to extract its global geometric character at a glance. Since the work by Becker and Karplus [6], the *disconnectivity graph* (DG) has been widely used to visualize an overall topography of the PESs. A PES is partitioned into basins of attraction of local minima (LMs) corresponding to stable configurations. The boundaries of basins are characterized by saddle points (SPs), which energetically control the connectivity between the adjacent LMs. The DG is a topological reconstruction which represents how a single metabasin is decomposed into basins by decreasing the total energy. Hence, in the DG, the number of energetically accessible metabasins and the LMs contained in each metabasin are encoded as functions of the total energy. These indeed provide necessary information to estimate the static properties [6,14]. However, the DG does not give information about all possible pathways on the PES. Unfortunately, the information about the number of hops over SPs necessary to travel from one LM to another, not depicted in the DG, is indispensable to evaluate the dynamic properties of the PES.

In this paper, we would like to develop a unified approach for understanding PES topography and dynamic properties. One of the most promising ways is a direct comparison of the detailed topography of a PES and the dynamics on the PES. To this end, we propose a topological reconstruction representing the network of individual saddle connectivity in an explicit way, which we call the *connectivity graph* (CG). Its availability is demonstrated by applying it to various PESs. By analyzing the transition frequencies between the LMs on them, we show that CGs play a crucial role in describing the nonequilibrium dynamics.

We start by describing how the CG is constructed from a given PES. Let the PES have LM i with energy E_i^{LM} $(i=1,2,\ldots,M)$ and SP j with E_j^{SP} $(j=1,2,\ldots,N)$, where the energies are sorted in ascending order. In the CG, LMs are represented by half vertical lines from E_i^{LM} to ∞. To include the individual connectivity, SP j, which connects a pair of LMs, is indicated by a horizontal segment attached to the corresponding vertical lines of the pair. Then, the lines of LMs are placed at even intervals in the order $\mathrm{LM}_{\sigma(1)},\mathrm{LM}_{\sigma(2)},\ldots,\mathrm{LM}_{\sigma(M)}$, where the permutation σ is designed not to lose the information about the disconnectivity. We specify σ by using a metabasin analysis as follows. Initialize the ordered set R as $\{\{1\},\{2\},\ldots,\{M\}\}$ and repeat the following step from $i=1$ to N: if SP i connects two elements of R_a and R_b $(a<b)$, then unite them by appending elements of R_b to R_a. The resulting R has only one element R_1 and σ is given by $\sigma=R_1$. Note that σ is normalized such that the global minimum (GM) of LM 1 is placed on the extreme left, i.e., $\sigma(1)=1$. Due to the construction, the CG describes the disconnectivity of the PES. More important is that, as we had expected, the CG exposes the direct connection between adjacent minima by a single horizontal segment connecting two vertical lines drawn to represent these minima.

In emphasizing the difference between DGs and CGs, we classify saddles into two types, *disconnective saddles* (DSs) and *nondisconnective saddles* (NDSs). When the total energy of the system is lowered below the potential energy of a certain DS, the basin that is connected by the DS splits up into two basins. In contrast to this, in the vicinity of a NDS energy, there is no decomposition of the basin resulting from the interruption of the NDS. Thus, the DSs represent the energetically cheapest pathways to travel. The NDSs are more expensive but can be dynamically relevant pathways between some two minima. It should be noted that the CG has the capability to represent both the DSs and the NDSs, while the DG is unable to represent the NDSs [15,16].

For illustration, we introduce a two-dimensional random funnel model (RFM). This model is generated by the superposition of bivariate Gaussian potentials, located on integer points (n_x,n_y) with $n_x,n_y=1,2,\ldots,L$, with an equal standard deviation $\tilde{\sigma}$ and random amplitudes from the uniform distribution on the interval from $V(n_x,n_y)$ to $V(n_x,n_y)+1$. In addition, the periodic boundary condition is imposed on the unit square $(0,L]\times(0,L]$ [17]. In order to put a single-funnel character in the present model, we employ a funnel-shaped function $V(x,y)=2F\max\{|x-L/2|,|y-L/2|\}/L$ of funnel depth F. The left panels of Fig. 1 show contour plots of RFMs for various F with an identical seed of the random

*okushima@ike-dyn.ritsumei.ac.jp

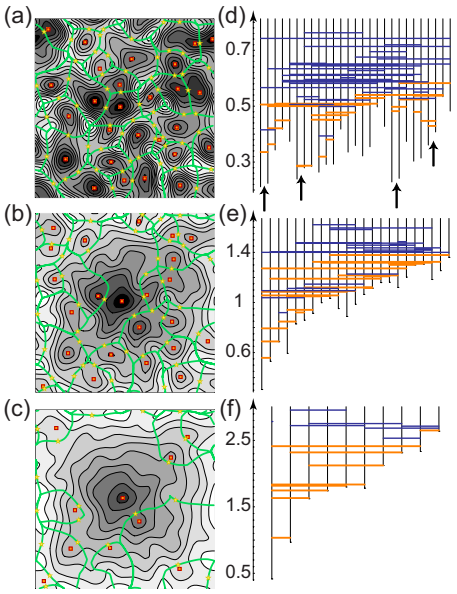

FIG. 1. (Color online) Left panels: Contour plots of RFMs with $L=16$ and $\tilde{\sigma}=0.84$ for $F=0$, 1, and 2.4 in (a), (b), and (c), respectively, where the LMs (boxes), the SPs (stars), and the boundaries of the LM basins [green (gray) bold lines] are depicted. Right panels: The corresponding CGs are presented in (d), (e), and (f) with the vertical axes of potential energy: the orange (gray) and blue (black) lines denote the pathways that run through DSs and NDSs, respectively. In (d), funnel bottoms are indicated by arrows.

number generator. These figures show that a larger F yields a smoother funnel-shaped PES. The configuration space of the RFM of $F=0$ in Fig. 1(a) is divided into multiple LM basins of approximately equal areas. When F becomes 1, however, the funnel character becomes apparent as shown in Fig. 1(b), where the GM near the center $(x,y)=(8,8)$ sinks down and its basin becomes dominating. Figure 1(c) for $F=2.4$ shows that a larger F yields a smoother funnel-shaped PES, where the GM becomes deeper and some LMs at $F=1$ disappear.

Let us examine how these RFMs are visualized with the CGs. The right panels of Fig. 1 are the CGs corresponding to the left contour plots, respectively. From Fig. 1(d) of the CG with $F=0$, we see that the PES is composed of five funnels. It further shows that the random-energy LM lines are randomly connected, at various saddle energy levels, with $(M-1)$ DS lines in a relatively low energy region $E^{SP} \lesssim 0.6$, as well as with many extra NDS lines in a relatively higher region, $E^{SP} \gtrsim 0.5$, where some NDSs are coexistent with DSs in the range $0.4 \lesssim E^{SP} \lesssim 0.6$. The CG with $F=1$ in Fig. 1(e) exhibits a rugged single-funnel feature clearly, where many direct links that connect the GM to the surrounding LMs appear close together. In this case, too, many NDS lines connect a lot of pairs of LMs both in the coexistent and in the higher regions. A typical CG for a smooth funnel-shaped PES is plotted in Fig. 1(f) for $F=2.4$, where all direct links to the GM from the adjacent LMs become DS lines, meaning that for smoother funnels the DSs provide better approximate descriptions of the low-energy topography. Nevertheless, NDSs, and thus CGs, are indispensable for valid topographic descriptions of generic PESs (see below). In short, CGs correctly represent the networks between LMs connected by NDSs, as well as DSs, and further show the smoothness and steepness of PESs as a funnel shape.

Now we proceed to test how CGs capture dynamic properties. For this purpose, an average passage time from a LM a to another LM b, i.e., rate constant, is computed as a dynamic indicator. According to the flux over population method [18], an external sink and source are connected to the domains of LM b and LM a, respectively, to build a nonequilibrium steady-state current J, and then the rate of the passage is given by $k=J/\rho$, where ρ is the population outside the domain of the sink. In what follows, we consider a simple LM basin hopping dynamics as a reference to statistically normal dynamics [6,14]. The intra-LM dynamics loses its memory so fast as to forget the previous inter-LM hoppings, and the inter-LM dynamics is represented by coupled master equations: $dP_i(t)/dt=\Sigma_{j\neq i}[k_{ij}P_j(t)-k_{ji}P_i(t)]$, where $P_i(t)$ denotes the occupation probability in LM i at time t and k_{ji} is the rate constant for a direct transition from LM i to an adjacent LM j. In the transition state theory, k_{ji} is given, at inverse temperature $\beta=(k_{\mathrm{B}}T)^{-1}$, by $k_{ji}=(1/\beta h) \times [Z_j^{\dagger}(\beta)/Z_i(\beta)]e^{-\beta(E_j^{SP}-E_i^{LM})}$, where SP j is the saddle standing in the way of transition from LM i to LM j, $Z_j^{\dagger}(\beta)$ is the transition state partition function of SP j, and $Z_i(\beta)$ is the partition function of the basin of minimum LM i [14].

We qualitatively know that at low temperatures the energetically favorable DS contribution dominates the hopping dynamics, while at high temperatures the energetically expensive NDS contribution may become superior due to its fewer number of steps. Then the actual question is how and why the most relevant pathways that carry most dominant probability current are switched on complicated, multidimensional PESs by changing the temperature [19]. For a systematic search for these superior pathways, we use the following formula [12,16,20,21], for the steady-state rate constant along an n-step pathway ℓ $(i_n \equiv b \leftarrow i_{n-1} \leftarrow \cdots \leftarrow i_2 \leftarrow i_1 \leftarrow i_0 \equiv a)$:

$$\frac{1}{k_\ell} = e^{-\beta E_{i_0}^{LM}} Z_{i_0}(\beta) \sum_{m=1}^{n} e^{\beta E_{i_m}^{SP}} / Z_{i_m}^{\dagger}(\beta). \tag{1}$$

The form of Eq. (1) enables us to consistently set $\exp(\beta E_{i_m}^{SP})/Z_{i_m}^{\dagger}(\beta)$ as the positive definite cost of the elementary transition $(i_m \leftarrow i_{m-1})$. Thus, the most relevant path, which contributes most dominantly to $k_{b\leftarrow a}$, is characterized by the most inexpensive one that minimizes this routing cost with the specified boundary conditions of starting at LM a and ending in LM b. Then the most inexpensive paths of interest are evaluated methodically by using the well-known Dijkstra shortest-path algorithm [16], which provides minimum spanning pathways coming to a single sink node LM b from the others [22,23].

Here, we apply the shortest-path algorithm to the RFMs. The CGs in Fig. 2 draw the shortest pathways coming into LM 1 from the other LMs for the RFM of $F=1$ for $\beta=0$, 5.7, 72, and 73, respectively. The CG (a) shows shortest pathways in the high-temperature limit. At high energy, many

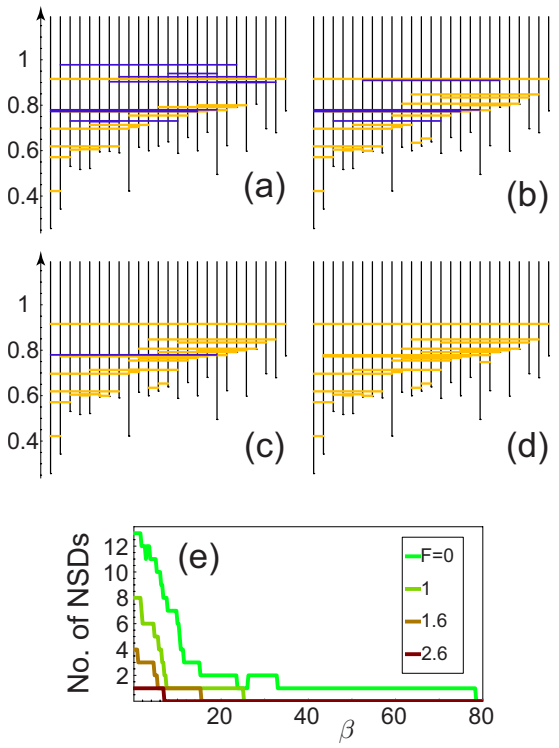

FIG. 2. (Color online) Dynamical shortest paths for RFM of $F=0.4$ are represented in a CG for (a) $\beta=0$, (b) 5.7, (c) 72, and (d) 73 [24]. (e) The numbers of NDSs on the shortest pathways vs β, for various F (inset).

energetically expensive but very shortcut NDSs contribute to the shortest pathways. For larger β, however, the energy is not so high that some energetically cheaper shortcut NDSs that are coexistent with DSs in the energy domain become used as shown in Fig. 2(b). The most competing NDS with smallest energetic penalty remains on the shortest pathways, up to a temperature $\beta \sim 72$ [Fig. 2(c)]. For the lower temperatures, all shortest paths are composed of all DSs, because no energetic penalty is allowed any longer as a member of the shortest paths [Fig. 2(d)]. To see how the switches of the shortest paths depend on temperature more closely, we plot the numbers of NDSs belonging to the shortest paths as functions of β for various F in Fig. 2(e). We see that for all RFMs the numbers of NDSs gradually decrease as the temperature is decreased. Furthermore, smoother funnel PESs cause more rapid decreases of NDSs. Note that, for $F=0$, a few NDSs remain very competitive even at rather low temperatures $\beta \sim 80$, because their energetic penalties are overcome with their shortness.

Finally, we show the applicability to a realistic multidimensional system, the 38-atom Lennard-Jones cluster (LJ_{38}). The well-known difficulty in visualization arises from the huge number of LMs [14]. Here, the monotonic sequence scheme with low barriers [14,25], which identifies the LMs in a funnel as a single LM [the so-called "monotonic sequence basin" (MSB)] is adopted as a coarse-graining method for CG. By iteratively applying this scheme, one obtains a series of coarse-grained CGs, namely, $CG_1, CG_2, \ldots, CG_n, \ldots, CG_{n_{\max}}$, describing the connection

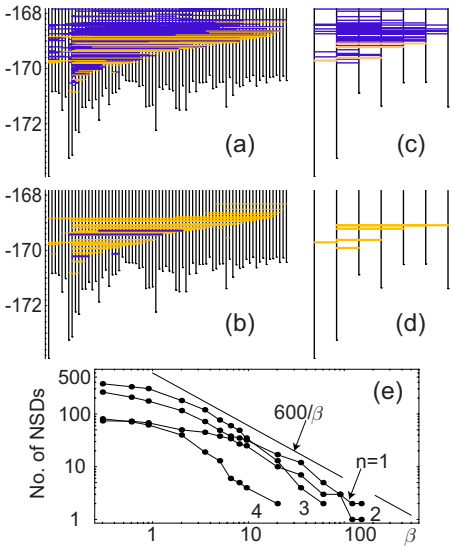

FIG. 3. (Color online) Dynamical shortest paths for LJ_{38} [24] are represented in the second-order coarse-grained CG at $\beta=$(a) 1, (b) 35, and in the third order at $\beta=$(c) 1 and (d) 35. (e) The numbers of NDSs, $N(n-1)-N(n)$, belonging only to the nth-order coarse-graining level on the shortest pathways plotted as functions of β (see text).

among MSBs at more macroscopic levels specified by n. Figure 3 shows the dominant pathways determined by the shortest path algorithm for $\beta=1$ and 35, in the second-order coarse-grained CG_2 [Figs. 3(a) and 3(b)] and in the third-order CG_3 [Figs. 3(c) and 3(d)]. Figures 3(a) and 3(c) show that, at high temperature, a great many NDS are on the dominant pathways. The left-end vertical MSB lines are loosely connected by horizontal shortest saddle lines, which shows that the MSB holding the GM is decoupled from other MSBs. The above observation just agrees with the fact that the GM of LJ_{38} is not easily accessed by relaxation dynamics [13]. Even at low temperature, relatively many NDSs remain shortest [Fig. 3(b)], because the multidimensional competition allows a rich variety of energetically degenerative dynamical paths, as expected from the results of the RFM. In contrast, in macroscopic description of the third-order CG_3 [Fig. 3(d)], we see no shortest NDSs. This observation shows that, at low temperature, nondisconnective shortest paths disappear from the higher-order CG, being absorbed into lower-order funnels. To elucidate this separation process, the number of the shortest NDSs belonging only to the nth-order coarse-graining level, namely, $N(n-1)-N(n)$ $(1 \leqslant n \leqslant n_{\max})$ where $N(n)$ is the number of NDSs in the nth-order CG representation ($n_{\max}=4$, for this case), is shown in Fig. 3(e) as a function of β. We see that, as the temperature is lowered, the numbers of macroscopic NDSs $(n=3,4)$ decrease fast and eventually reach zero, respectively, at around $\beta=60,20$, while those of microscopic NDSs $(n=1,2)$ decrease slowly. Furthermore, the number of NDSs in each level approximately depends on β as $\propto \beta^{-1}$ in the scaling regime $\beta \gtrsim 2$ (the straight line in the figure is $600/\beta$ to guide the eye), indicating a self-similar structure among saddle connections in various levels n.

In summary, we have developed a topographic mapping of PESs, called a CG, which is an extended DG to include dynamically relevant NDSs. With this graph, we have confirmed that the dynamically relevant pathways are frequently switched on a temperature change, by studying the kinetics of Markov basin hopping on several PES models. With the help of Dijkstra's shortest-path algorithm, we developed an efficient method to systematically compute all the most dominant pathways. We computed the numbers of NDSs on the shortest pathways as an indicator of the multiplicity in the connection on the PES and elucidated its relation to funnel smoothness with a simple two-dimensional random funnel model. The usefulness of our method is also demonstrated with a realistic Lennard-Jones cluster. These results show that the CG provides us a sound basis for understanding the multiple connectivity in dynamics and structures on complex many-dimensional PESs.

ACKNOWLEDGMENTS

The authors acknowledge and express their gratitude to D. J. Wales for permitting the use of LJ_{38} data for the minima and transition states [19]. This work was supported by Grants-in-Aid for Scientific Research on Priority Areas (2) Nos. 14077213 and 14077220 from the Ministry of Education, Culture, Sports, Science and Technology of Japan. The authors are grateful to S. Tsuji for his hospitality.

[1] J. P. K. Doye and L. Meyer, Phys. Rev. Lett. **95**, 063401 (2005).
[2] R. S. Berry and R. Breitengraser-Kunz, Phys. Rev. Lett. **74**, 3951 (1995).
[3] J. P. K. Doye, Phys. Rev. Lett. **88**, 238701 (2002).
[4] M. Menon, A. N. Andriotis, D. Strivastava, I. Ponomareva, and L. A. Chernozatonskii, Phys. Rev. Lett. **91**, 145501 (2003); Y.-H. Kim, I. H. Lee, K. J. Chang, and S. Lee, *ibid.* **90**, 065501 (2003) and references cited therein.
[5] S. V. Krivov, S. F. Chekmarev, and M. Karplus, Phys. Rev. Lett. **88**, 038101 (2002).
[6] O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).
[7] H. W. Sheng *et al.*, Nature (London) **439**, 419 (2006).
[8] P. G. Debenedetti and F. H. Stillinger, Nature (London) **410**, 259 (2001); F. H. Stillinger, Science **267**, 1935 (1995).
[9] M. Winterlich, G. Diezemann, H. Zimmerman, and R. Bohmer, Phys. Rev. Lett. **91**, 235504 (2003).
[10] B. Doliwa and A. Heuer, Phys. Rev. Lett. **91**, 235501 (2003).
[11] A. Heuer, Phys. Rev. Lett. **78**, 4051 (1997).
[12] B. Doliwa and A. Heuer, Phys. Rev. E **67**, 031506 (2003).
[13] D. J. Wales *et al.*, Nature (London) **394**, 758 (1998); D. J. Wales and H. A. Scheraga, Science **285**, 1368 (1999).
[14] For a recent review, see D. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, U. K. 2004).
[15] Moreover, it is easy to extend the CG to impose information about higher-order saddles.
[16] Full details will be presented elsewhere.
[17] The potential is low-pass filtered with a cutoff wave number π for computational convenience.
[18] P. Hänggi *et al.*, Rev. Mod. Phys. **62**, 251 (1990), and reference therein.
[19] D. J. Wales, Mol. Phys. **100**, 3285 (2002); **102**, 891 (2004).
[20] H. Eyring and E. M. Eyring, *Modern Chemical Kinetics* (Reinhold, New York, 1963).
[21] We suppose that only a unique path is selected because of exponential dominance [16].
[22] E. W. Dijkstra, Numer. Math. **1**, 269 (1959).
[23] For applications to discrete path sampling methods for PESs, see D. J. Wales, J. Chem. Phys. **121**, 1080 (2004); **122**, 234903 (2005).
[24] Here all frequency factors are set to 1 for simplicity, and the saddle lines in CGs are colored as in Fig. 1.
[25] S. Saito *et al.*, Adv. Class. Traj. Meth. **4**, 105 (1999).