



Title	Effect of timescale on energy landscape: Distinction between free-energy landscape and potential of mean force
Author(s)	Kawai, Shinnosuke; Komatsuzaki, Tamiki
Citation	Physical Review E, 87(3), 030803 <a href="https://doi.org/10.1103/PhysRevE.87.030803">https://doi.org/10.1103/PhysRevE.87.030803</a>
Issue Date	2013-03-14
Doc URL	<a href="http://hdl.handle.net/2115/52651">http://hdl.handle.net/2115/52651</a>
Rights	© 2013 American Physical Society
Type	article
File Information	PhysRevE.87.030803.pdf



[Instructions for use](#)

# Effect of timescale on energy landscape: Distinction between free-energy landscape and potential of mean force

Shinnosuke Kawai\* and Tamiki Komatsuzaki

*Molecule & Life Nonlinear Sciences Laboratory, Research Institute for Electronic Science, Hokkaido University, Kita 20 Nishi 10, Kita-ku, Sapporo 001-0020, Japan*

(Received 11 October 2012; published 14 March 2013)

We investigate the effects of the timescale of motion on the shape of energy landscapes. The distinction between the free-energy landscape and the potential of mean force is clarified. The former is related to a thermal equilibrium distribution for chosen coordinates, whereas the latter is determined by the mean force exerted on the coordinates in the equilibrium. It is found that the condition for these two energy landscapes to be the same is the constancy of the mean square velocity with respect to the position coordinate. However, even when the condition holds for the chosen coordinates, as the timescale of observation increases, the averaging effect causes a decrease in the mean square velocity nonuniformly in the configuration space, resulting in a larger distinction between the free-energy landscape and the potential of mean force. The results are expected to provide an important basis for the study of long time scale phenomena in large systems.

DOI: [10.1103/PhysRevE.87.030803](https://doi.org/10.1103/PhysRevE.87.030803)

PACS number(s): 82.20.-w, 02.50.-r, 05.45.-a, 82.37.-j

Energy landscape, the energy of a given system expressed as a function of the coordinates, is one of the most versatile concepts in the study of multidimensional systems. Its application can be found in the studies of, for example, large molecules, such as clusters [1,2] and proteins [3–10], diffusion in oxide networks [11], fullerene materials [12], formation process of gels and glasses [13], chemical and biological separation through micro- and nanofluidic systems [14], and epigenetics in developmental biology [15–17]. Two important roles of the energy landscape can be pointed out. One is that the gradients of the energy landscape with respect to the coordinates give the forces in the directions of those coordinates by which the motion of the system is driven on the landscape. Thus, chemical reactions can be imagined as a point mass moving on the energy surface drawn in the configuration space, and their rate constants can also be calculated from the information of the energy surface [8,18–21]. Also, phenotypic diversification of cells can be modeled as the motion of marbles on Waddington's landscape [15–17].

The other role of the energy landscape is that the probability distribution at the stationary state is given by the energy landscape through the Boltzmann factor. Thus, the stable structures of proteins can be predicted as minima of the free-energy landscape, and it is possible to elucidate specific interactions that give rise to the stable structures [7,22–24]. It is, in a sense, surprising that these two different concepts, one being a dynamical concept of force and the other being a statistical distribution, can be given by the same energy function.

Large complex systems involve a huge number of degrees of freedom and experience the underlying energy landscapes with hierarchical time and space scales. When some, but not all, degrees of freedom are considered to have reached a stationary distribution, it is natural to represent the energy landscape with a smaller set of degrees of freedom by averaging all the others over the stationary distribution. Moreover, due to

limited experimental time resolution or an interest in slow motions, e.g., relevant to biological functions, rather than the full details of motion over all the timescales, we are often interested in investigating the behavior of physical quantities over long timescales. What kinds of energy landscape are seen at different timescales in such reduced systems is an intriguing subject. It is even not clear whether the two different roles mentioned above can still be played by a unique energy surface. In this Rapid Communication, we discuss in what circumstances these two different roles can or cannot be played by a single energy landscape and investigate how the timescale of observation affects energy landscapes to provide insights for understanding the long-term dynamics of large systems.

Corresponding to the two roles, there exist two kinds of energy landscapes: One is the potential of mean force [18,25], and the other is the free-energy landscape [3–6]. Let us now formulate the situation in mathematical terms. Suppose  $\mathbf{X} = (X_1, X_2, \dots, X_N)$  is the collection of all the coordinates required to describe the system. An example may be the set of the position coordinates and conjugate momenta of all the particles in a classical many-particle system. In thermal equilibrium, the coordinate  $\mathbf{X}$  is distributed according to the equilibrium distribution  $\rho_{\text{eq}}(\mathbf{X})$ . We can choose some function  $Q(\mathbf{X})$  to represent the system (e.g., the number of native contacts or the largest principal components in a protein) and average out all the rest. ( $Q$  can be a collection of functions, but, for the sake of simplicity, we consider here only one function.) Let us denote by the uppercase  $Q$  the function of  $\mathbf{X}$  and its particular value by the lowercase  $q$ . Then, the equilibrium distribution for the value of  $Q$  is given by

$$P(q) = \int \delta[Q(\mathbf{X}) - q] \rho_{\text{eq}}(\mathbf{X}) d\mathbf{X}, \quad (1)$$

where  $\delta$  denotes the Dirac  $\delta$  function. Then, the *free-energy landscape*  $G(q)$  can be defined as such a function that gives the equilibrium distribution as the Boltzmann factor,

$$\exp\left(-\frac{G(q)}{k_B T}\right) \propto P(q), \quad (2)$$

\*skawai@es.hokudai.ac.jp

with some appropriately defined temperature. The free-energy landscape is a coarse-grained description of the system since its definition includes averaging over the bath modes (all the degrees of freedom other than  $Q$ ) as in Eq. (1). It is different from the *potential-energy surface*  $V(X_1, X_2, \dots, X_{N/2})$ , which is defined on each point in the configuration space (if we take  $(X_1, \dots, X_{N/2})$  as the position coordinates and  $(X_{N/2+1}, \dots, X_N)$  as their conjugate momenta).

On the other hand, the *mean force*  $f_M(q)$  can be defined as the average of the acceleration for each fixed value  $q$  [25],

$$f_M(q) = \langle \ddot{Q}; q \rangle = \frac{1}{P(q)} \int \ddot{Q}(X) \delta[Q(X) - q] \rho_{\text{eq}}(X) dX, \quad (3)$$

where  $\ddot{Q}(X)$  is the second-order time derivative of  $Q(X)$ , which is also a function of  $X$ , and the notation  $\langle \ddot{Q}; q \rangle$  stands for the equilibrium average of  $\ddot{Q}$  with  $Q$  fixed at the value  $q$ . Then, the *potential of mean force*  $V_{\text{MF}}(q)$  is defined as such a function that the mean force is given by its gradient,

$$V_{\text{MF}}(q) = - \int^q f_M(q') dq'. \quad (4)$$

As terminology, we note that the definitions stated here are not unequivocal ones. In the literature, these terms, free-energy landscape and potential of mean force, have sometimes been used with the same meaning. Efforts to rationalize the free-energy landscape of proteins at the single molecule level [26,27] and in the context of Anfinsen's dogma [28] have been carried out. However, to avoid confusion, we tentatively adopt the definitions of Eqs. (2) and (4) in this Rapid Communication. It is far from trivial whether these two landscapes,  $G(q)$  and  $V_{\text{MF}}(q)$ , are equal to each other or not.

Pope and Ching [29], Ching [30], and Stolovitzky and Ching [31] found the following expression for the stationary distribution of general stochastic systems:

$$P(q) = \langle \delta[Q(X) - q] \rangle \propto \frac{1}{\langle \dot{Q}^2; q \rangle} \exp \left( \int^q \frac{\langle \ddot{Q}; q' \rangle}{\langle \dot{Q}^2; q' \rangle} dq' \right), \quad (5)$$

where  $\langle \cdot \rangle$  is the ensemble average taken over the stationary distribution. This equation can be proved generally for any stationary stochastic process where the stochastic variable is twice differentiable and  $P(q)$  decreases rapidly as  $|q|$  goes to infinity.

By comparing Eq. (5) with Eqs. (2)–(4), one can see that the free-energy landscape  $G(q)$  and the potential of mean force  $V_{\text{MF}}(q)$  become the same if

$$\langle \dot{Q}^2; q \rangle = k_B T \quad (\forall q), \quad (6)$$

in other words, if the mean square of the velocity does not depend on  $q$  and is interpreted as “temperature.”

Let us now look into the effect of timescale of observation on these energy landscapes. The issue of the timescale and the energy landscape has been stated by several authors [26,32,33]. It was pointed out that large molecules, such as proteins, explore different morphologies of the energy landscape for different timescales of motion. In the present Rapid Communication, we investigate the change in the energy landscapes defined above with coarse graining of the

observation time. It will be found that, even when the condition Eq. (6) holds for the original coordinate  $q$ , these landscapes become different as the observation timescale is increased and local averaging is performed.

Suppose our “observable” is a local time average of the original variable  $Q$ , and let us denote it by  $\bar{Q}$ ,

$$\bar{Q}(t) = \frac{1}{\tau} \int_{t-\tau/2}^{t+\tau/2} Q(t') dt', \quad (7)$$

with a specified time window  $\tau$ . Then, we can think of the energy landscapes for the new variable  $\bar{Q}$ , which are not necessarily the same as those for  $Q$ .

We demonstrate the change in the energy landscape with the change in  $\tau$  by a simple one-dimensional system given by a Langevin equation,

$$\ddot{Q} = - \frac{\partial}{\partial Q} V_0(Q) - \gamma \dot{Q} + \xi(t), \quad (8)$$

where the Morse potential is adopted as the potential  $V_0$ ,

$$V_0(Q) = D_e \{1 - \exp[-\alpha(Q - Q_e)]\}^2, \quad (9)$$

with  $D_e = 1$ ,  $\alpha = 1$ , and  $Q_e = 0$ . The friction constant is taken as  $\gamma = 0.1$ . The random force  $\xi(t)$  is a Gaussian white noise whose autocorrelation is given by the fluctuation-dissipation relation,

$$\langle \xi(t) \xi(t') \rangle = 2k_B T \gamma \delta(t - t'), \quad (10)$$

with  $k_B T = 0.2$ . The Morse potential is depicted in Fig. 1(a). The Morse potential is steeper in the region  $q < 0$  and is flatter in  $q > 0$ . Consequently, the motion is expected to typically take place more rapidly in  $q < 0$  and more slowly in  $q > 0$ .

Figure 1 shows the results of calculations of the free-energy landscape, the potential of mean force, and the mean square velocity for the original coordinate  $Q$  and the locally averaged quantity  $\bar{Q}$  with the time window  $\tau = 1$  and 2. For averaging, the numerical simulation of Eq. (8) was performed for 4000 sample trajectories of the time length of 1000. The simulation has been performed by the method of Ref. [34] with a time step of  $10^{-4}$ . The free-energy landscape is then calculated from the distribution according to Eq. (2) with the effective temperature defined by  $k_B T_{\text{eff}} = \langle (d\bar{Q}/dt)^2 \rangle$ . The potential of mean force is calculated by Eq. (4) with the mean force Eq. (3).

Figure 1(a) shows the potential energy  $V_0(q)$  and the mean square velocity  $\langle \dot{Q}^2; q \rangle$  as functions of  $q$  before taking the local average. In this case, the mean square velocity does not depend on  $q$  and is always equal to  $k_B T = 0.2$ . Consequently, as Eqs. (5) and (6) tell us, the free energy and the potential of the mean force are the same and are the same as the original potential  $V_0$  as shown in Fig. 1(a).

In contrast, Fig. 1(b) shows the results for the locally averaged quantity  $\bar{Q}$  as defined in Eq. (7) with the time window of  $\tau = 1$ . The mean square velocity  $\langle \dot{\bar{Q}}^2; \bar{q} \rangle$  is now lower than  $k_B T = 0.2$  since the local time averaging cancels out fast changes in the velocity. This effect of averaging out the velocity fluctuations is more pronounced at smaller values of  $q$ . Figure 2 shows an example trajectory in this system to observe the effect of local time averaging. When the coordinate  $q$  takes large values (e.g.,  $t \approx 380$ –387 in Fig. 2), the motion is relatively slow due to the mild slope of the potential. Then,

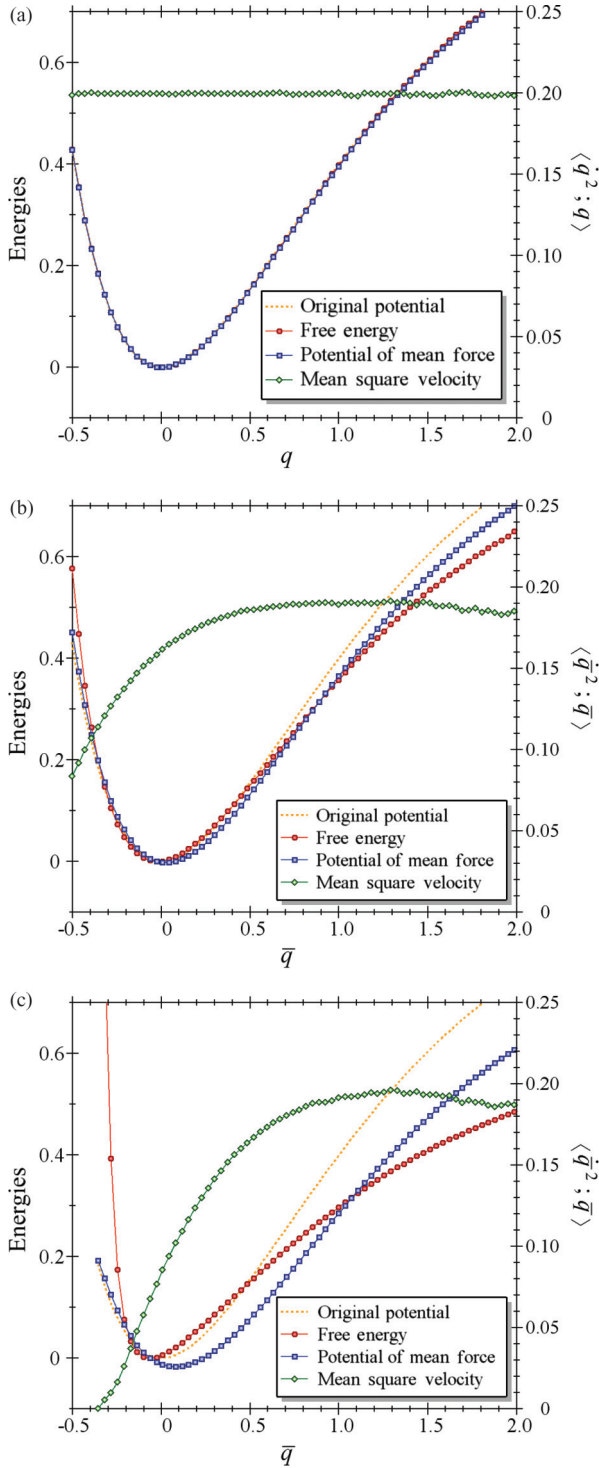


FIG. 1. (Color online) Dotted curve, left axis: The Morse potential  $V_0(q)$  used in the present calculation; circles, left axis: the free-energy landscape  $G(\bar{q})$ ; squares, left axis: the potential of mean force  $V_{MF}(\bar{q})$ ; and diamonds, right axis: the mean square velocity  $\langle \dot{q}^2; \bar{q} \rangle$ . (a) As functions of the original coordinate  $q$ . (b) For the local time average  $\bar{q}$  of  $q$  with the time window of  $\tau = 1$ . (c) With the time window of  $\tau = 2$ . Scaled units are used so that the two parameters of the Morse potential and the mass of the system are scaled to one.

the time-averaged trajectory follows the original trajectory to a good approximation, and therefore, its velocity is closer to that of the original trajectory. On the other hand, when the

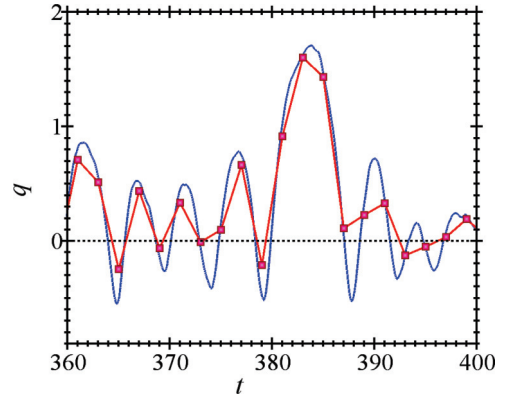


FIG. 2. (Color online) Thick solid curve: Example trajectory for the Morse potential with friction and random force. The local-average trajectory  $\bar{q}$  with a time window size of  $\tau = 1$  is shown by squares. Scaled units are used so that the two parameters of the Morse potential and the mass of the system are scaled to one.

coordinate  $q$  takes small values (e.g.,  $t \approx 364\text{--}366$ ,  $369\text{--}370$ ,  $373\text{--}375$ ,  $378\text{--}380$ , and so forth, in Fig. 2), the system is strongly kicked back in a short time by the steep slope of the potential. The averaged trajectory cannot trace this fast bouncing-back motion, and its velocity becomes small due to the averaging of the forward and backward movements contained within one time window. This explains the fact that, with the same window size  $\tau$ , the effect of taking the time average is more significant for the smaller  $q$  than for the larger  $q$  region. The consideration made around Eq. (6) tells that the free energy and the potential of mean force are not necessarily the same if the mean square velocity is not uniform. Figure 1(b) shows that this is indeed the case as there is a small discrepancy between these two.

The nonuniformity of the mean square velocity is more pronounced in Fig. 1(c) where we take a larger time window of  $\tau = 2$ . As a consequence, the discrepancy between the free-energy landscape and the potential of mean force have become larger. As a trend, the free-energy landscape rises more steeply as  $\bar{q}$  decreases than the potential of mean force, whereas an opposite behavior [ $G(\bar{q})$  flatter than  $V_{MF}(\bar{q})$ ] is observed for positive  $q$ . We can physically interpret this behavior as follows: For a given slope of  $V_{MF}(\bar{q})$ , it is “more difficult” to climb up the potential surface if the system has smaller velocity. As  $V_{MF}$  increases with the change in  $\bar{q}$ , therefore, the population decreases more significantly for smaller  $\langle \dot{Q}^2; \bar{q} \rangle$  than for larger  $\langle \dot{Q}^2; \bar{q} \rangle$ . Thus, the free-energy landscape  $G(\bar{q})$ , being calculated from the distribution, increases more sharply than the potential of mean force  $V_{MF}(\bar{q})$  when the mean square velocity  $\langle \dot{Q}^2; \bar{q} \rangle$  is smaller than the average  $\langle \dot{Q}^2 \rangle$ . It is noted that such a discrepancy between the potential of mean force and the free-energy landscape can happen in general, despite the simplicity of the model system we have chosen here. The discrepancy between these two landscapes arises from the inherent property of the system concerning the comparison between the timescale of the observation and that of the system motion in each region of the potential, influenced by conditions such as the morphology of the potential surface, friction, and temperature.

To summarize, we have shown that the two kinds of energy landscape, the free-energy landscape and the potential of mean force, can be significantly different. The condition for them to coincide is the uniformity of the mean square velocity. We also demonstrated the effect of time coarse graining on the landscapes. As the time window becomes large, the mean square velocity decreases by local averaging. When the system has different timescales of motion in different regions of the configuration space, the decrease in the mean square velocity is not uniform, giving rise to the nonuniformity of the mean square velocity, and therefore, the free-energy landscape and the potential of mean force become different. Since interest in studying large scale without being involved in the details of fast motions is very common in physical research, the results of this Rapid Communication, although demonstrated with a

simple model system, will provide an important basis for the study of slow motions of large systems.

We are grateful to Professor M. Toda from Nara Women's University and Professor S. Nara from Okayama University for fruitful discussions. This work has been partially supported by the Human Frontiers Science Program Grant No. RGP0031/2010, a Grant-in-Aid for Scientific Research(B), JSPS, a Grant-in-Aid for Young Scientists(B), JSPS, a Grant-in-Aid for Exploratory Research, JSPS, and a Grant-in-Aid for Scientific Research on Innovative Areas "Spying minority in biological phenomena (Grant No. 3306)," MEXT. The computations were partially performed using the Research Center for Computational Science, Okazaki, Japan.

- 
- [1] K. D. Ball, R. S. Berry, R. E. Kunz, F.-Y. Li, A. Proykova, and D. J. Wales, *Science* **271**, 963 (1996).
  - [2] D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, UK, 2003).
  - [3] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, *Science* **254**, 1598 (1991).
  - [4] J. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, *Proteins: Struct., Funct., Genet.* **21**, 167 (1995).
  - [5] P. G. Wolynes, *Philos. Trans. R. Soc. London, Ser. A* **363**, 453 (2005).
  - [6] D. J. Wales, *Curr. Opin. Struct. Biol.* **20**, 3 (2009).
  - [7] B. Tarus, J. E. Straub, and D. Thirumalai, *J. Mol. Biol.* **379**, 815 (2008).
  - [8] Y. Suzuki and O. K. Dudko, *Phys. Rev. Lett.* **104**, 048101 (2010).
  - [9] O. K. Dudko, T. G. W. Graham, and R. B. Best, *Phys. Rev. Lett.* **107**, 208301 (2011).
  - [10] G. G. Maisuradze, A. Liwo, and H. A. Scheraga, *Phys. Rev. Lett.* **102**, 238102 (2009).
  - [11] A. Bongiorno and A. Pasquarello, *Phys. Rev. Lett.* **88**, 125901 (2002).
  - [12] S. De, A. Willand, M. Amsler, P. Pochet, L. Genovese, and S. Goedecker, *Phys. Rev. Lett.* **106**, 225502 (2011).
  - [13] S. Jabbari-Farouji, G. H. Wegdam, and D. Bonn, *Phys. Rev. Lett.* **99**, 065701 (2007).
  - [14] J. A. Bernate and G. Drazer, *Phys. Rev. Lett.* **108**, 214501 (2012).
  - [15] J. Wang, L. Xu, and E. Wang, *Proc. Natl. Acad. Sci. USA* **105**, 12271 (2008).
  - [16] R. Sekine, M. Yamamura, S. Ayukawa, K. Ishimatsu, S. Akama, M. Takinoue, M. Hagiya, and D. Kiga, *Proc. Natl. Acad. Sci. USA* **108**, 17969 (2011).
  - [17] C. H. Waddington, *The Strategy of the Genes* (Allen and Unwin, London, 1957).
  - [18] R. D. Levine, *Molecular Reaction Dynamics* (Cambridge University Press, Cambridge, UK, 2005).
  - [19] S. L. Mielke, K. A. Peterson, D. W. Schwenke, B. C. Garrett, D. G. Truhlar, J. V. Michael, M.-C. Su, and J. W. Sutherland, *Phys. Rev. Lett.* **91**, 063201 (2003).
  - [20] J. Zhang, D. Dai, C. C. Wang, S. A. Harich, X. Wang, X. Yang, M. Gustafsson, and R. T. Skodje, *Phys. Rev. Lett.* **96**, 093201 (2006).
  - [21] W. Shiu, J. J. Lin, and K. Liu, *Phys. Rev. Lett.* **92**, 103201 (2004).
  - [22] B. Kuhlman and D. Baker, *Curr. Opin. Struct. Biol.* **14**, 89 (2004).
  - [23] Y. Zhang, *Curr. Opin. Struct. Biol.* **18**, 342 (2008).
  - [24] R. Das, B. Qian, S. Raman, R. Vernon, J. Thompson, P. Bradley, S. Khare, M. D. Tyka, D. Bhat, D. Chivian *et al.*, *Proteins: Struct., Funct., Bioinf.* **69**, 118 (2007).
  - [25] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, London, 2001).
  - [26] A. Baba and T. Komatsuzaki, *Proc. Natl. Acad. Sci. USA* **104**, 19297 (2007).
  - [27] T. Komatsuzaki, A. Baba, S. Kawai, M. Toda, J. Straub, and R. Berry, *Adv. Chem. Phys.* **145**, 171 (2011).
  - [28] A. Ben-Naim, *Chem. Phys. Lett.* **511**, 126 (2011).
  - [29] S. B. Pope and E. S. C. Ching, *Phys. Fluids A* **5**, 1529 (1993).
  - [30] E. S. C. Ching, *Phys. Rev. Lett.* **70**, 283 (1993).
  - [31] G. Stolovitzky and E. S. Ching, *Phys. Lett. A* **255**, 11 (1999).
  - [32] H. Yang, G. Luo, P. Karnchanaphanurach, T.-M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie, *Science* **302**, 262 (2003).
  - [33] J. E. Straub and D. Thirumalai, *Proc. Natl. Acad. Sci. USA* **90**, 809 (1993).
  - [34] D. L. Ermak and H. Buckholz, *J. Comput. Phys.* **35**, 169 (1980).