

# Dynamical study of metallic clusters using the statistical method of time series clustering

S.K. Lai<sup>a,b,\*</sup>, Yu-Ting Lin<sup>a</sup>, P.J. Hsu<sup>a,b</sup>, S.A. Cheong<sup>c</sup>

<sup>a</sup> Complex Liquids Laboratory, Department of Physics, National Central University, Chungli 320, Taiwan

<sup>b</sup> Molecular Science and Technology Program, Taiwan International Graduate Program, Academia Sinica, Taipei 115, Taiwan

<sup>c</sup> Division of Physics and Applied Physics, School of Physics and Mathematical Sciences, Nanyang Technological University, SPMS-04-01, 21 Nanyang Link, Singapore 637371, Singapore

## ARTICLE INFO

### Article history:

Received 14 September 2009

Received in revised form 24 October 2010

Accepted 29 December 2010

Available online 5 January 2011

### Keywords:

Molecular dynamics simulation

Time series data analysis

Metallic clusters

## ABSTRACT

We perform common neighbor analysis on the long-time series data generated by isothermal Brownian-type molecular dynamics simulations to study the thermal and dynamical properties of metallic clusters. In our common neighbor analysis, we introduce the common neighbor label (CNL) which is a group of atoms of a smaller size (than the cluster) designated by four numeric digits. The CNL thus describes topologically smaller size atomic configurations and is associated an abundance value which is the number of “degenerate” four digits all of which characterize the same CNL. When the cluster is in its lowest energy state, it has a fixed number of CNLs and hence abundances. At nonzero temperatures, the cluster undergoes different kinds of atomic activities such as vibrations, migrational relocation, permutational and topological isomer transitions, etc. depending on its lowest energy structure. As a result, the abundances of CNLs at zero temperature will change and new CNLs with their respective new abundances are created. To understand the temperature dependence of the CNL dynamics, and hence shed light on the cluster dynamics itself, we employ a novel method of statistical time series analysis. In this method, we perform statistical clustering at two time scales. First, we examine, at given temperature, the signs of abundance changes at a short-time scale, and assign CNLs to two short-time clusters. Quasi-periodic features can be seen in the time evolution of these short-time clusters, based on which we choose a long-time scale to compute the long-time correlations between CNL pairs. We then exploit the separation of correlation levels seen in these long-time correlations to extract strongly-correlated collections of CNLs, which we will identify as effective variables for the long-time cluster dynamics. It is found that certain effective variables show subtleties in their temperature dependences and these thermal traits bear a delicate relation to prepeaks and main peaks seen in clusters  $\text{Ag}_{14}$ ,  $\text{Cu}_{14}$  and  $\text{Cu}_{13}\text{Au}_1$ . We therefore infer from the temperature changes of effective variables and locate the temperatures at which these prepeaks and principal peaks appear, and they are evaluated by comparing with those deduced from the specific heat data.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

At very low temperatures, atoms in a metallic cluster execute solid-like vibration. The description of this oscillatory motion in the presence of an external probe such as temperature is quite different from that in the bulk. Whereas in a bulk system the thermal response of atoms are treated individually the same owing to the translational symmetry, atoms in a cluster under the same condition are, however, realized by their whereabouts locations since their geometrical sites have much bearing on how they respond and hence the cluster properties. Consider, for example, a 14-atom pure cluster and a 38-atom bimetallic cluster

$\text{Ag}_{32}\text{Cu}_6$  at their respective lowest energy states. The former [1,2] can be differentiated generally into three broad kinds of atoms, namely, floating (adatom), surface and central atoms and the latter [3], on the other hand, can be categorized coarsely into surface and central atoms. Such a classification of atoms by the positions they reside is instructive at higher temperatures and provides a practical means in quantitative analysis of the microscopic dynamics of clusters. In two recent communications [2,4], we have applied this strategy of partitioning atoms in a cluster to investigate the velocity autocorrelation function; we deduced from this dynamic quantity and its Fourier-transformed function, the power spectrum, the average temperature at which a cluster melts. We demonstrated that these dynamic variables can be employed to infer the melting temperature  $T_m$  of a cluster and the predicted value is reasonably close to that estimated from the principal peak position of the specific heat  $C_V$  (widely accepted to be the melt-

\* Corresponding author at: Complex Liquids Laboratory, Department of Physics, National Central University, Chungli 320, Taiwan.

E-mail address: sklai@coll.phy.ncu.edu.tw (S.K. Lai).

ing temperature  $T_m$ ). In the course of interpreting the dynamic results, an attempt was made to explain the anomalous increase of the relative root-mean-square bond length fluctuation constant

[2,4]  $\delta = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i}^n \frac{\sqrt{(r_{ij}^2(t))_t - (r_{ij}(t))_t^2}}{(r_{ij}(t))_t}$  whose aberrant thermal response occurs not at  $T_m$  but at a much lower temperature. In these latter works [2,4], attention was drawn to tracking down simultaneously the instantaneous relative bond length  $r_{ij}(t)$  in  $\delta$  for  $ij$ -atoms. A careful and thorough analysis of  $r_{ij}(t)$  at each temperature reveal indeed the complex dynamics of individual atoms. It is now understood, for instance, that the anomalous increase of  $\delta$  at lower temperatures for  $\text{Cu}_{14}$  or  $\text{Ag}_{14}$  [2,4] is due mainly to the migrational relocations of the floating atom. As the temperature of the cluster climbs up further, the  $r_{ij}(t)$  indicates an increase in the frequency of permutational isomer transitions among the floating and surface atoms, and such exchange activities continue until the central atom participates finding its way out via permutating with a surface or floating atom. The cluster should by then approach  $T_m$ . Although the analysis using  $r_{ij}$  is dynamically fruitful, it is nevertheless a complicated procedure since a synchronous follow-up of different  $r_{ij}$ s (atomic pairs of floating-surface, surface-surface, central-surface, etc.) must be effected for a conclusive description to be reached. Moreover, the  $r_{ij}$  variable does not give much insight into the factors that govern the atomic dynamics at different temperatures.

In this work, we turn to a more practical means of studying the dynamical behavior of clusters. Rather than examining  $r_{ij}$  simultaneously for different pairs of  $ij$ -atoms, we scrutinize instead a group of atoms which comprises a root pair and its common neighbors. This kind of the common neighbor analysis (CNA) finds its application in a wide range of problems in bulk systems. We illustrate in this paper that the CNA can equally well be applied to study the dynamical motion of atoms in a cluster. Differing from the usual simulation method where the time development of the position coordinates and velocities of atoms in a cluster are used to explore the cluster dynamics from calculating such quantities as the root-mean-square displacement, velocity autocorrelation function, etc., we introduce here a new scheme – based on CNA and time series clustering – to understand the time evolution.

In a typical molecular dynamics (MD) simulation running over tens of millions of time steps, the sheer volume of data contained in the large number of displacement time series makes analyzing and understanding the underlying physics a challenging task. Simple visualization of these simulation data in the form of a movie brings forth too much subjective bias. An objective way to reduce data complexity is to effect time series clustering, a class of methods that has found applications in diverse areas of business, science and technology. In time series clustering, the objective is to group together time series data sets with similar dynamical features so that a much smaller collection of dynamically distinct clusters is obtained. Broadly, the time series clustering methods can be classified into model-based methods [5–11] and correlations-based methods [12–18]. Both these methods have been reviewed by Liao [19] and we refer the interested readers to this article for more details. Since the physics behind nanocluster melting is not yet fully understood, we cannot apply existing model-based methods to our problem, as this would introduce modeling biases. Neither can we simply adapt existing correlations-based methods, as our goal is to develop a mechanistic picture of the melting process. Instead, we choose to develop a new correlations-based time series clustering method that is a hybridization of the methods developed by Rummel et al. [14,18] and Tumminello et al. [13,15]. In our method, we draw attention to novel ways of analyzing the correlation matrix and demonstrate, by illustration of several small clusters using CNA, how the *effective variables* (to be defined below) are obtained and applied

to describe the dynamics of our physical system. We should emphasize that the time series clustering method was developed for larger complex systems in mind, and has in fact been successfully applied to finance [20,21], neuroscience [22–26], and meteorology [27–29]. We have, for instance, very recently applied the method to financial markets (500–3000 degrees of freedom) [30] and global positioning system networks (100 degrees of freedom) [31]. In these two systems, the dynamical time scales are not well separated. Employing the method of partial hierarchical clustering [30], we, however, managed to extract the effective variables as well as their dynamics. In contrast, the dynamical time scales are well separated in the metallic clusters chosen here, up to very high temperatures. This important qualitative difference between the financial/geophysical systems and metallic clusters confers an added advantage to the time series clustering method, because the separation of dynamical time scales translates into a separation of correlation levels in each of the long-time windows. There is thus no cause for alarm in the event that the dynamical time scales are not well separated. Armed with this time series clustering scheme, we proceed to interpret the microscopic time series data which we will obtain from an isothermal Brownian-type MD simulation on several metallic clusters.

The present paper is organized as follows. In Section 2, we give a brief summary of the CNA following the definition of Honeycutt and Andersen [32] and introduce the method of time series clustering on multiple time scales in detail. We devote a significant portion of the method to explain its advantage. In the discussions below, we refer extensively to various statistical observations made from the time series of the common neighbor labels which are sole elements in the CNA. Our intention is to explain clearly to readers how quasi-regular synchronies in the short-time dynamics can be converted into reliable long-time correlations, based on which we then identify the effective variables. We shall not describe the isothermal Brownian-type MD simulations that were used to generate the microscopic time series data. We refer the interested readers, however, to our previous works for technical details [1,2,4]. Throughout this section, for concreteness, we use  $\text{Cu}_{13}$  as an illustrative example to inaugurate the statistical time series clustering methodology. In Section 3, we give a quantitative discussion of the results derived from the time series analysis first for  $\text{Ag}_{14}$ , and then for  $\text{Cu}_{14}$  and  $\text{Cu}_{13}\text{Au}_1$ ; we make a comparison between  $\text{Cu}_{14}$  and  $\text{Cu}_{13}\text{Au}_1$ , delving into the possible interpretations of the dynamics. With regard to the melting scenario, we infer  $T_m$  from the temperature-dependences of the time series data and compare its value with corresponding one deduced from  $C_V$ . Since the  $C_V$  for  $\text{Ag}_{14}$ ,  $\text{Cu}_{14}$  and  $\text{Cu}_{13}\text{Au}_1$  have been reported previously [1,2], we will therefore simply cite the results or just summarize them without further description. Finally, we give a conclusion in Section 4.

## 2. Methodology: Common neighbor analysis and time series analysis

### 2.1. Common neighbor analysis

In our CNA, we examine the neighborhoods around all pairs of atoms in the cluster, and note these at each time step in the form of four-digit common neighbor labels (CNL),  $\{c_1, c_2, c_3, c_4\}$ . In this analysis, we say that there is a *bonded* root-pair atoms if the distance between them is  $r_b \leq 1.2r_0$ ,  $r_0$  being the nearest neighbor separation. Therefore, for a given *root-pair* atoms, we set  $c_1 = 1$  if they are bonded, or  $c_1 = 2$  if they are not. Next, we set  $c_2$  equal to the number of atoms bonded simultaneously to both root-pair atoms. We call these the common neighbors of the root-pair atoms. Then, we set  $c_3$  equal to the number of bonds among the common-neighbor atoms. Finally, we can have

topologically-distinct neighborhoods with CNL that have exactly the same first three digits. These are distinguished by the fourth digit  $c_4 = 1, 2, \dots$  in their CNLs. As examples, we illustrate how two such CNLs are determined for the  $\text{Cu}_{13}$  cluster in Fig. 1.

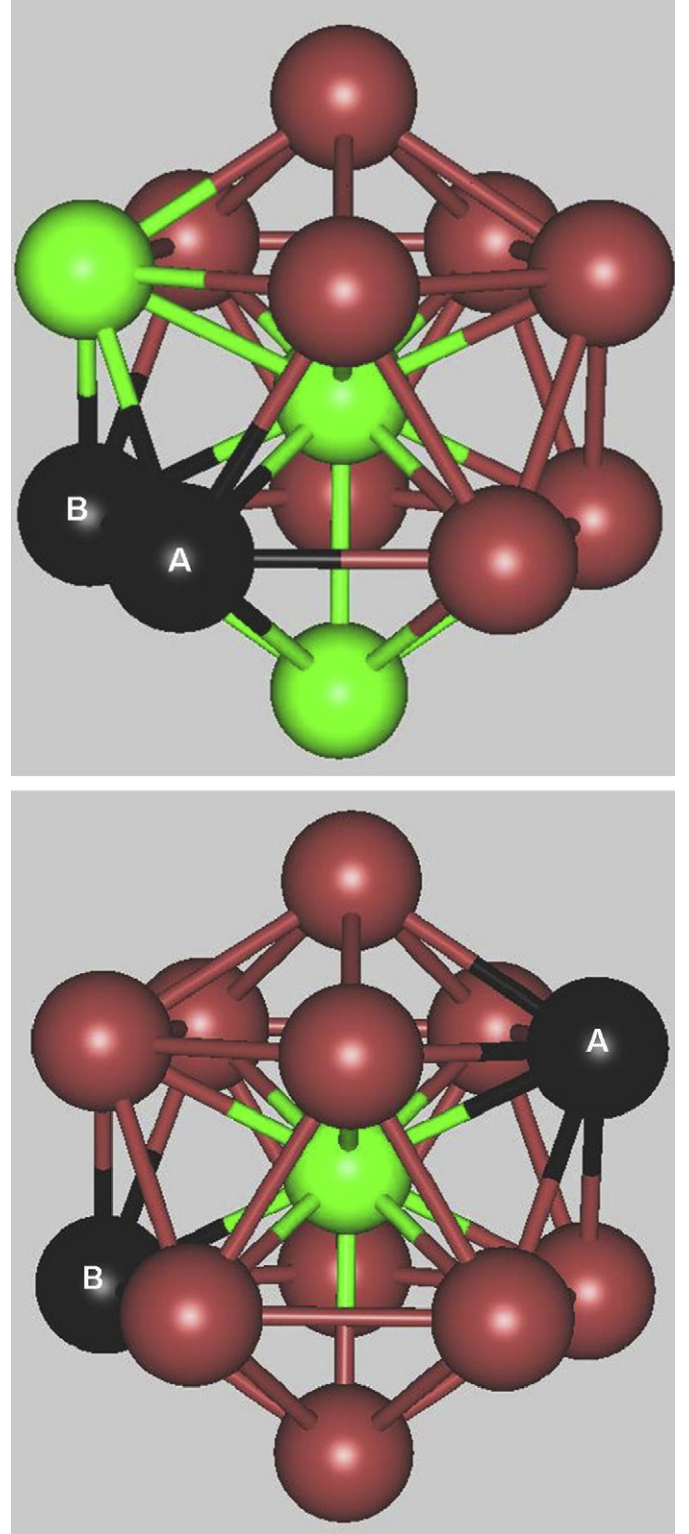
## 2.2. Time series analysis

In this section, we describe the general framework for extracting effective variables using the statistical method of time series clustering at multiple time scales. Effective variables are collections of microscopic variables which evolve in time synchronously. These are identified in two stages. First, we examine the time series data over a sequence of non-overlapping short-time windows. Within each short-time window, microscopic variables are assigned to various short-time clusters based on a chosen short-time statistic. For concreteness, let us consider the CNLs {1321} and {2101} of  $\text{Cu}_{13}$  shown in Fig. 1, in an isothermal Brownian-type MD simulation running over  $1 \times 10^7 \Delta t$  (each time step  $\Delta t = 0.001$  ps). We choose a short-time window of  $\delta_{sw} = 10\Delta t$ . At  $T = 0$  K, the CNLs {1321} and {2101} have, respectively, a total of thirty and six “degenerate”  $\{c_1, c_2, c_3, c_4\}$  and these are called their *abundances*. For  $T > 0$ , these abundances change from time to time, as a result of atomic motions within the cluster. We identify the abundance  $p_{i,k}$  of the  $i$ th CNL at the  $k$ th short-time window as the microscopic variable, and define the short-time statistic by the CNL correlation level

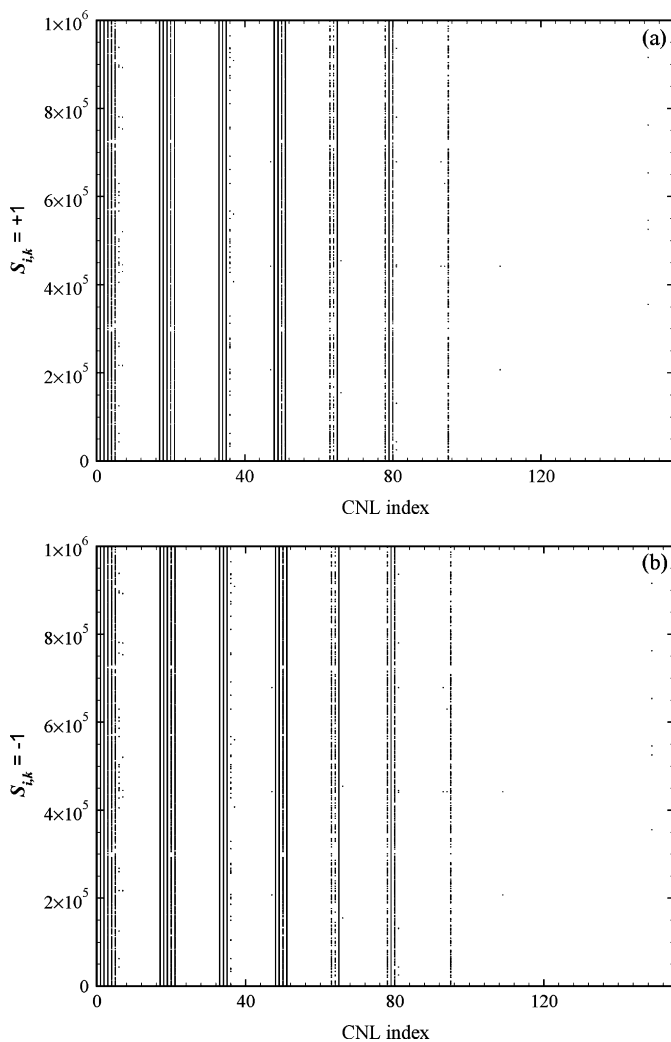
$$S_{i,k} = \begin{cases} +1, & \Delta p_{i,k} > 0; \\ 0, & \Delta p_{i,k} = 0; \\ -1, & \Delta p_{i,k} < 0, \end{cases} \quad (1)$$

where  $\Delta p_{i,k} = p_{i,k+1} - p_{i,k}$ . Based on the values of this short-time statistic  $S_{i,k}$ , we can then assign microscopic variables to three different short-time clusters at each short-time step  $k$ . Of these, only the  $S_{i,k} = +1$  and  $S_{i,k} = -1$  short-time clusters provide dynamical information on the behavior of the cluster. A simple way to visualize the dynamics of the short-time statistic  $S_{i,k}$  over long times is to make a density plot, whereby the  $(i, k)$  cell is filled if  $S_{i,k} = +1$  (or  $S_{i,k} = -1$  if we are interested in the other dynamically-informative short-time cluster). In Fig. 2(a)–(b), we show such density plots for the short-time clusters  $S_{i,k} = \pm 1$  of  $\text{Cu}_{13}$  at  $T = 800$  K. In these figures, we label the CNLs by indices. A look-up table is constructed to track down the CNL by referring to the CNL index (see Table 1). Returning to Fig. 2(a)–(b), it is in general not easy to discern regularities in the density plots when we show the entire duration of the simulation. These regularities become clearer when we restrict ourselves to a shorter time duration (but still long compared to  $\delta_{sw}$ ), as shown in Fig. 3(a)–(b) for  $T = 800$  K. We see quasi-periodic features in the density plots that involve more CNLs and vary at a shorter period (compare, for example, to that at  $T = 400$  K (not shown)). Also, we have studied the density plots obtained using different short-time windows,  $\delta_{sw} = 10$  and  $50\Delta t$ . We find that a quasi-periodic feature with a period between  $1000$  to  $3000\delta_{sw}$  can be seen more clearly for  $\delta_{sw} = 10\Delta t$ . This tells us that the dynamical information captured by our short-time statistic  $S_{i,k}$  is robust with respect to our choice of this short-time window  $\delta_{sw}$ . This robustness is important, because we can then select a long-time window  $\delta_{LW}$  based on the unambiguous observed period and carry out a coarse-grained analysis in the second stage of our effective variable extraction framework. To effect the long-time dynamics, we construct within each long-time window a long-time correlation matrix  $\mathbf{C}$ , whose matrix elements are defined by

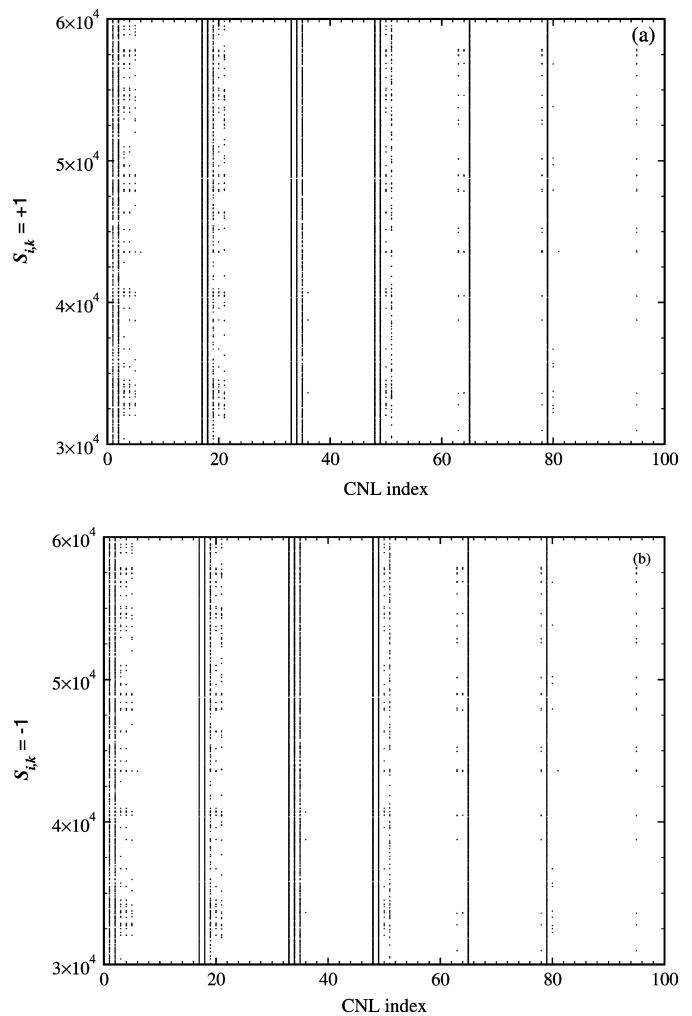
$$C_{ij} = (1 - \delta_{ij}) \sum_{k=1}^m \frac{1}{2} |S_{i,k} S_{j,k} (S_{i,k} + S_{j,k})| \quad (2)$$



**Fig. 1.** The CNLs {1321} (top) and {2101} (bottom) for the  $\text{Cu}_{13}$  cluster. In this figure, atoms are shown as solid spheres, while bonds between atoms are shown as connecting rods. The root-pair atoms are colored black, and labeled A and B, while their common-neighbor atoms are colored green. In the top subfigure, we find that A and B are bonded root-pair atoms, and share three common neighbors. This is the first topologically-distinct neighborhood where two bonds are found among the three common-neighbor atoms. Hence the CNL is {1321}. In the bottom subfigure, we find that A and B are not bonded root-pair atoms, but they share a common neighbor. Since there is only one common-neighbor atom, there is no bond within the common neighborhood of A and B. Hence we denote the CNL by {2101}. (For interpretation of the references to colors in this figure, the reader is referred to the web version of this article.)

**Table 1**Look-up table for CNL and its corresponding index for Cu<sub>13</sub> at 800 K.

CNL index	CNL			CNL index	CNL		
1	1	1	0	1	48	2	3
2	2	1	0	1	49	1	4
3	1	2	0	1	50	2	4
4	2	2	0	1	51	1	5
5	1	3	0	1	63	1	4
6	2	3	0	1	64	2	4
7	1	4	0	1	65	1	5
17	1	2	1	1	66	2	5
18	2	2	1	1	78	2	4
19	1	3	1	1	79	1	5
20	2	3	1	1	80	2	5
21	1	4	1	1	81	1	6
33	1	3	2	1	93	1	5
34	2	3	2	1	94	2	5
35	1	4	2	1	95	1	6
36	2	4	2	1	109	1	6
37	1	5	2	1	149	1	5
47	1	3	3	1			

**Fig. 3.** Same as Fig. 2 but for selected shorter time durations.

of  $\mathbf{C}$  are distinguished by their matrix elements  $C_{ij}$  significantly small. Based on this decomposition of  $\mathbf{C}$ , we identify the collections of strongly-correlated CNLs associated with such distinct submatrices as our effective variables.

As a concrete illustration of this long-time clustering procedure to identify effective variables, we perform a Brownian-type MD run for Cu<sub>13</sub> at  $T = 700$  K. For a total number of  $1 \times 10^7 \Delta t$  simulation

**Fig. 2.** CNL correlation levels (a)  $S_{i,k} = +1$  and (b)  $S_{i,k} = -1$  of Cu<sub>13</sub> calculated at  $T = 800$  K for short-time clusters.

where  $m$  is the number of short-time window steps within the long-time window. Note that the long-time correlation matrix  $\mathbf{C}$  is symmetric, i.e.  $C_{ij} = C_{ji}$ , and the value of the matrix element  $C_{ij}$  is the number of times the CNL correlation levels  $S_{i,k}$  and  $S_{j,k}$  are assigned to the same short-time cluster within the chosen long-time window. The diagonal elements  $C_{ii}$  are set to zero, because they provide no cross-correlational information on the cluster dynamics. The long-time correlation matrix  $\mathbf{C}$ , thus contains matrix elements of different correlation levels. These are the product of dynamics operating on different time scales (such as the collection of time scales shown in Fig. 4). Dynamics at time scales much shorter than the long-time window will self-average, producing a low background correlation level contribution to  $\mathbf{C}$ . Dynamics at time scales longer than the long-time window, on the other hand, will produce correlation levels rising above the background. Even without knowing the precise nature of the dynamical structures, we expect dynamics with time scales closest to the long-time scale to produce stronger correlation levels, and dynamics with time scales further from the long-time scale to produce weaker correlation levels. If these dynamical time scales are widely separated, there will be a corresponding clean separation of correlation levels in the long-time correlation matrix  $\mathbf{C}$ . Where this separation of correlation levels is found, we can directly decompose  $\mathbf{C}$  into distinct submatrices, whereby the large matrix elements  $C_{ij}$  of  $\mathbf{C}$  are found only within these submatrices, and any other submatrices



time steps, we choose the long-time window to be  $\delta_{LW} = 2000\delta_{SW}$ . Advancing the long-time window at intervals of  $1000\delta_{SW}$ , we obtain 999 *overlapping* long-time windows each of which is associated with a  $\mathbf{C}$  matrix. In Fig. 5(a), we give the long-time correlation matrix of the first long-time window, which runs from  $\delta_{SW}$  to  $2000\delta_{SW}$ . As we can see, there are several distinct groups of matrix

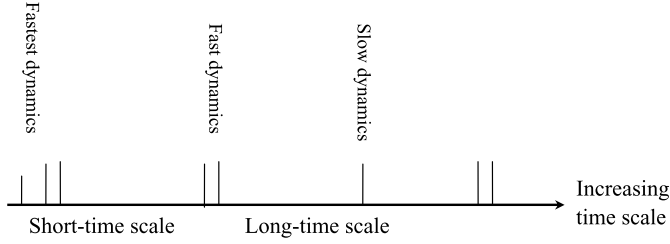


Fig. 4. Schematic diagram showing different time scales.

elements:  $C_{ij} > 200$ ,  $C_{ij} \sim 80$ ,  $C_{ij} \sim 60$ ,  $C_{ij} \sim 15$ , and  $C_{ij} < 10$ , representing distinct correlation levels within the long-time window. The frequency of occurrence of  $C_{ij}$  assuming the same numerical value in the matrix  $\mathbf{C}$  (see Fig. 5(a)) plotted against the magnitude of this numerical value of  $C_{ij}$  is depicted in Fig. 6. In particular, we find the  $C_{ij} \geq 200$  correlation level very strong, and are very well separated from the rest of the correlation levels. From the correlation matrix  $\mathbf{C}$ , one can extract the effective variables in terms of collections of CNLs (microscopic variables) which are grouped together on the basis of a cutoff value of  $C_{ij}$  (see Fig. 6).

Let us digress to describe how the effective variables are determined. Fig. 7 is the flow chart showing the numerical procedure that we follow to extract the effective variable. We start by first scrutinizing the histogram for the frequency of occurrence of the correlation level elements  $C_{ij}$  calculated for  $\text{Cu}_{13}$  at  $T = 700$  K. As indicated in Fig. 6, there is a large gap separating some high correlation level elements (whose  $C_{ij} \geq 200$ ) from other lower ones ( $C_{ij} \leq 82$ ). We then read off from Fig. 6 a cutoff value of

14	1101	2101	1211	2211	1311	1411	1321	2321	1421	2331	1431	1531	1541	1551
1101	0	56	56	59	9	4	0	57	52	0	4	5	1	0
2101	56	0	74	76	7	4	1	53	50	1	8	2	2	1
1211	56	74	0	281	6	3	2	264	80	2	213	11	203	0
2211	59	76	281	0	10	4	1	262	79	1	212	12	200	1
1311	9	7	6	10	0	1	0	6	3	0	2	5	0	0
1411	4	4	3	4	1	0	0	4	1	0	2	0	0	0
1321	0	1	2	1	0	0	0	20	1	289	9	1	9	285
2321	57	53	264	262	6	4	20	0	82	20	209	11	202	18
1421	52	50	80	79	3	1	1	82	0	1	29	2	28	1
2331	0	1	2	1	0	0	289	20	1	0	9	1	9	285
1431	4	8	213	212	2	2	9	209	29	9	0	8	208	7
1531	5	2	11	12	5	0	1	11	2	1	8	0	1	2
1541	1	2	203	200	0	0	9	202	28	9	208	1	0	5
1551	0	1	0	1	0	0	285	18	1	285	7	2	5	0

(a)

Fig. 5. (a) Matrix elements of correlation level matrix  $\mathbf{C}$  calculated for  $\text{Cu}_{13}$  at  $T = 700$  K. (b) Re-ordered matrix elements of the matrix  $\mathbf{C}$  given in (a). The submatrices shaded grey yield the same effective variables.

14	1211	1431	1541	2211	2321	1321	1551	2331	1101	2101	1421	1311	1411	1531
1211	0	213	203	281	264	2	0	2	56	74	80	6	3	11
1431	213	0	208	212	209	9	7	9	4	8	29	2	2	8
1541	203	208	0	200	202	9	5	9	1	2	28	0	0	1
2211	281	212	200	0	262	1	1	1	59	76	79	10	4	12
2321	264	209	202	262	0	20	18	20	57	53	82	6	4	11
1321	2	9	9	1	20	0	285	289	0	1	1	0	0	1
1551	0	7	5	1	18	285	0	285	0	1	1	0	0	2
2331	2	9	9	1	20	289	285	0	0	1	1	0	0	1
1101	56	4	1	59	57	0	0	0	0	56	52	9	4	5
2101	74	8	2	76	53	1	1	1	56	0	50	7	4	2
1421	80	29	28	79	82	0	0	0	52	50	0	3	1	2
1311	6	2	0	10	6	0	0	0	9	7	3	0	1	5
1411	3	2	0	4	4	1	1	1	4	4	1	1	0	0
1531	11	8	1	12	11	1	2	1	5	2	2	5	0	0

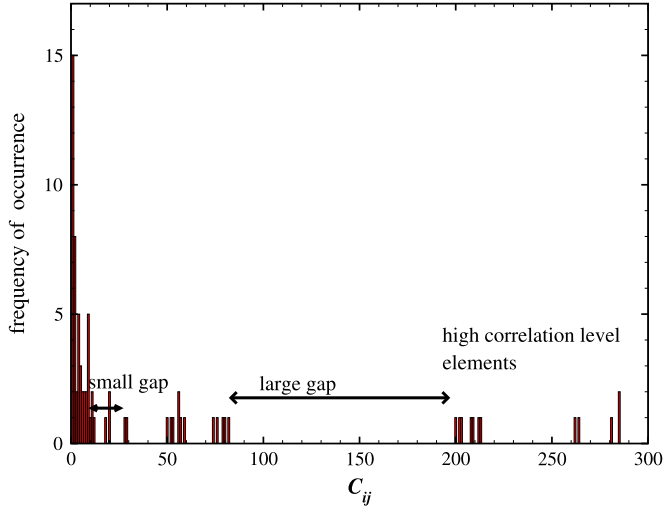
(b)

Fig. 5. Continued.

$C_{ij}$  which we denote as  $C_{cut}$ . Choosing  $C_{cut} = 141$ , we next turn to Fig. 5(a) again and read in column by column the elements  $C_{ij}$  of the correlation matrix  $\mathbf{C}$ . For the first column of  $C_{ij}$ ,  $i = 1$ , we check the inequality  $C_{ij} > C_{cut}$  for row elements  $j = 1, 2, \dots$ , of  $C_{1j}$ . If the answer is positive, we group these CNLs, otherwise we move on to the second column. For example, the first two columns in Fig. 5(a) produce no CNL that matches the inequality criterion. However, a group of five CNLs satisfies  $C_{3j} > C_{cut}$  for the 3rd column and they are {1211, 2211, 2321, 1431, 1541}. A further check for  $C_{ij} > C_{cut}$  among the (contracted group of) CNLs {2211, 2321, 1431, 1541} in this third column group is effected immediately. The result of this verification is that four CNLs remain. Next, we proceed to do similar checking among the three CNLs {2321, 1431, 1541} for  $C_{ij} > C_{cut}$ . This procedure is continued hierarchically, and we confirm that the 3rd column yields an effective variable {1211, 2211, 2321, 1431, 1541}. We should emphasize that this kind of hierarchical checking among CNLs has to be carried out in a manner until group(s) of CNLs are sorted out. The latter is/are then labeled as effective variable(s). For this first-window  $\mathbf{C}$ ,

the above hierarchical procedure is repeated for columns 4th, 7th, 8th, 10th, 11th, 13th and 14th (i.e. for  $i > 3$ ). A comparison of the effective variables obtained in all columns is finally effected to ensure that there is no overlapping effective variables. Following this numerical scheme, two effective variables emerge, namely, {1221, 1431, 1541, 2211, 2321} and {1321, 1551, 2331}. We should mention at this point that we can re-order the long-time correlation matrix displayed in Fig. 5(a) to that in Fig. 5(b). In this re-ordered  $\mathbf{C}$ , notice that the off-diagonal matrix elements within two submatrices are all greater than 200, and the matrix elements in these two submatrices are much larger than the matrix elements of any other submatrices whose  $C_{ij}$  are all significantly smaller than 200 (for example, {1101, 2101, 1421}). One can thus identify in this re-ordered long-time correlation matrix two diagonal blocks and they are in fact associated with the same effective variables. In general, when we slide our long-time window along the simulation period, the number of the effective variables can change, and so can CNL and number of CNLs within each effective variable. To examine the long-time variations of the effective variables, and

determine their dynamical time scales, as well as the nature and character of their long-time dynamics, we need to automate the extraction process. The key to this automation (shown in Fig. 7) is the separation of correlation levels, as shown in Fig. 6. Unlike existing methods [23–25,28], the present method is designed specifically for identifying effective variables from very large time series data sets in an automated and unbiased fashion, even when the natures and structures of the effective variables change slowly with time. By sliding the long-time window along the time axis, the method will also be able to identify dynamical time scales at which these effective variables evolve, provided that these effective time scales are significantly longer than the short-time scale.



**Fig. 6.** Frequency of occurrence of the element  $C_{ij}$  assuming the same numerical value in the correlation matrix  $\mathbf{C}$  of Fig. 5(a) for the first long-time window vs. the magnitude of this numerical value  $C_{ij}$  for  $\text{Cu}_{13}$  calculated at 700 K.

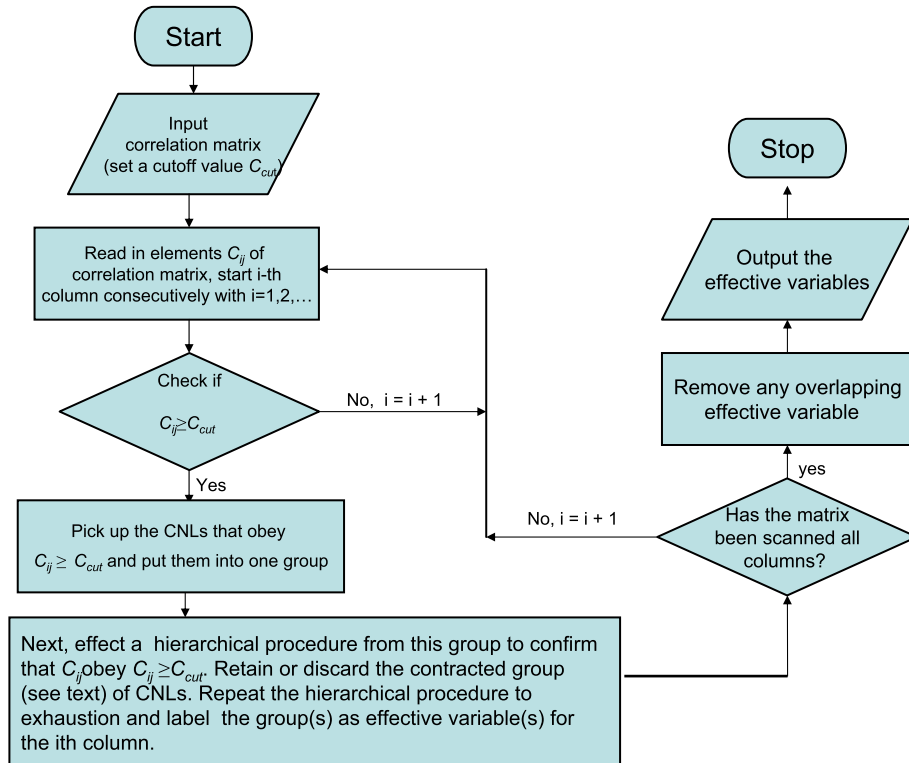
### 3. Results and discussions

We are now in a position to apply the time series analysis to study the dynamical properties of metallic clusters. We first present results for  $\text{Ag}_{14}$ , and then for  $\text{Cu}_{14}$  and  $\text{Cu}_{13}\text{Au}_1$ , in a manner that allows comparison between pure and bimetallic clusters.

#### 3.1. Pure cluster $\text{Ag}_{14}$

We have previously studied [2] the dynamical properties of  $\text{Ag}_{14}$ . The energy histogram and its lowest and lower energy structures are given in Ref. [2] to which the readers are referred for details. At temperatures  $T \leq 140$  K, this cluster is in its lowest energy state – an icosahedron with a floating atom or adatom residing near a three-atoms base. The first excited state (in addition to the ground state) occurs at  $T = 150$  K and it persists until near  $T = 200$  K at which temperature the second excited state emerges. Note that, apart from the distorted geometry of the icosahedron, the first and second excited states differ from the lowest energy state in that a deformed four-atoms base is located near the floating atom. At  $T = 300$  K, only the lowest, first and second excited states are found (see Fig. 2 in Ref. [2]). More excited states appear for  $T > 300$  K.

An instructive description of the long-time dynamics of the cluster  $\text{Ag}_{14}$  can be obtained by showing at each temperature how frequently the effective variable appears in all of the 999 long-time windows. Since the effective variable is associated with collections of CNLs, it is convenient in the following discussion to label the latter by an effective variable index (EVI). Table 2 is a look-up table for tracking down the effective variable. At given temperature, the occurrence of effective variables in all of the long-time windows are recorded and plotted against EVI. Let us begin by scrutinizing the density plots of the short-time clusters  $S_{i,k} = \pm 1$  obtained at  $T = 100$  K using  $\delta_{\text{sw}} = 10\Delta t$ . They are depicted in Fig. 8(a)–(b). Based on the quasi-period seen in the density plots,



**Fig. 7.** Flow chart showing the numerical procedure for extracting the effective variables.

**Table 2**Look-up table for the effective variables labeled by numeric integer index for Ag<sub>14</sub>.

Effective variable index	Number of CNL	CNLs							
1	8	1101	1201	1311	1541	2201	2211	2321	2431
2	5	1211	1321	1431	1551	2331			
3	4	1201	1311	2321	2431				
4	2	2201	2211						
5	3	1321	1551	2331					
6	2	1211	1431						
7	3	1541	2211	2321					
8	5	1211	1431	1541	2211	2321			
9	2	2211	2321						
10	2	1201	2211						
11	3	1201	2211	2321					
12	3	1101	2211	2321					
13	2	1201	1311						
14	6	1211	1431	1541	2101	2211	2321		
15	4	1101	2201	2321	2431				
16	3	1201	1311	2211					
17	3	1211	2211	2321					
18	2	1101	2201						
19	3	1201	1311	2431					
20	2	2101	2211						
21	2	1551	2331						
22	4	1321	1551	2101	2331				
23	2	1211	2321						
24	2	1321	2331						
25	3	1101	2201	2431					
26	2	1541	2211						
27	2	2201	2311						
28	2	1211	2211						
29	2	2321	2331						
30	2	1541	2321						
31	3	1211	1431	2211					
32	2	1211	2101						
33	4	1211	2101	2211	2321				
34	4	1211	1541	2211	2321				
35	2	1101	1311						
36	3	1211	2101	2211					
37	2	1431	2331						
38	2	1311	2211						
39	2	1101	2101						
40	3	1101	1201	2101					
41	2	1101	1201						
42	2	1201	2321						
43	2	1211	2331						
44	5	1101	1201	1311	2211	2321			
45	3	1101	2201	2311					
46	3	1101	2101	2201					
47	3	1211	1311	2211					
48	4	1211	1321	1431	2331				
49	3	1211	1421	2211					
50	3	1211	1321	2331					
51	3	1321	1431	2331					
52	4	1101	1201	1311	2211				
53	4	1101	1201	1311	2321				
54	3	1201	1311	2321					

we choose  $\delta_{LW} = 2000\delta_{SW}$ , and slide the long-time window in  $1000\delta_{SW}$  increments to generate 999 long-time windows, for a total simulation time steps of  $1 \times 10^7 \Delta t$ . At this temperature, we see in Fig. 9(a) only two effective variables, EVI-1 (hereafter we label the  $i$ th effective variable index by an integer  $i$  as EVI- $i$ ) and EVI-2. Checking the look-up Table 2, they are {1101, 1201, 1311, 1541, 2201, 2211, 2321, 2431} and {1211, 1321, 1431, 1551, 2331}, respectively. Since the energy histogram (see Fig. 2 in Ref. [2]) indicates that only the ground state is present at  $T = 100$  K, these two effective variables describe solely vibrational motion. Comparing the CNLs in the two effective variables with those in the lowest energy structure (Fig. 10), we see more CNLs of EVI-1 than the ground state structure and the EVI-1 accounts for the harmonic type oscillatory motion and captures atomic configurations with larger distortion in contrast to that displayed by EVI-2.

In Fig. 9(a)–(b), we show how the occurrence frequencies of effective variables depend on temperature, from  $T = 100$  to 1500 K, in intervals of 100 K. We see that as we increase the temperature ( $T > 100$  K), the number of effective variables increases rapidly to around 30 for  $T = 200$ –300 K after which it decreases rapidly to around 5 by  $T = 700$  K. More interestingly, we find that EVI-1 and EVI-2 appear more frequently as we go from  $T = 100$  to 200 K, but become less frequent at higher temperatures. As the occurrence frequencies of EVI-1 and EVI-2 decline, we notice a corresponding rise in the occurrence frequency of EVI-5, which, from Table 2, consists of the CNLs {1321, 1551, 2331}. The EVI-5 peaks at  $T = 400$  K, and thereafter decreases with increasing temperature, but it continues to be detected until at  $T = 900$  K where it disappears identically. While these dynamic behaviors unfold, EVI-24 whose CNLs are {1321, 2331} emerges at a temperature as low as  $T = 200$  K, and it keeps growing with increasing temperature. Since the CNL



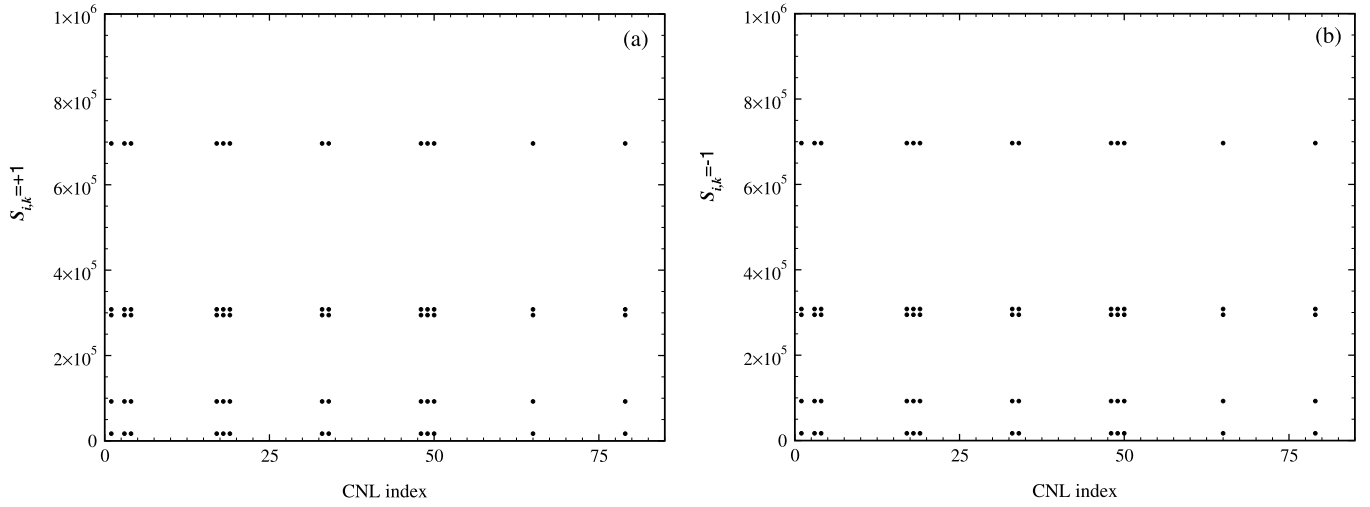


Fig. 8. CNL correlation levels (a)  $S_{i,k} = +1$  and (b)  $S_{i,k} = -1$  of  $\text{Ag}_{14}$  calculated at  $T = 100$  K for short-time clusters.

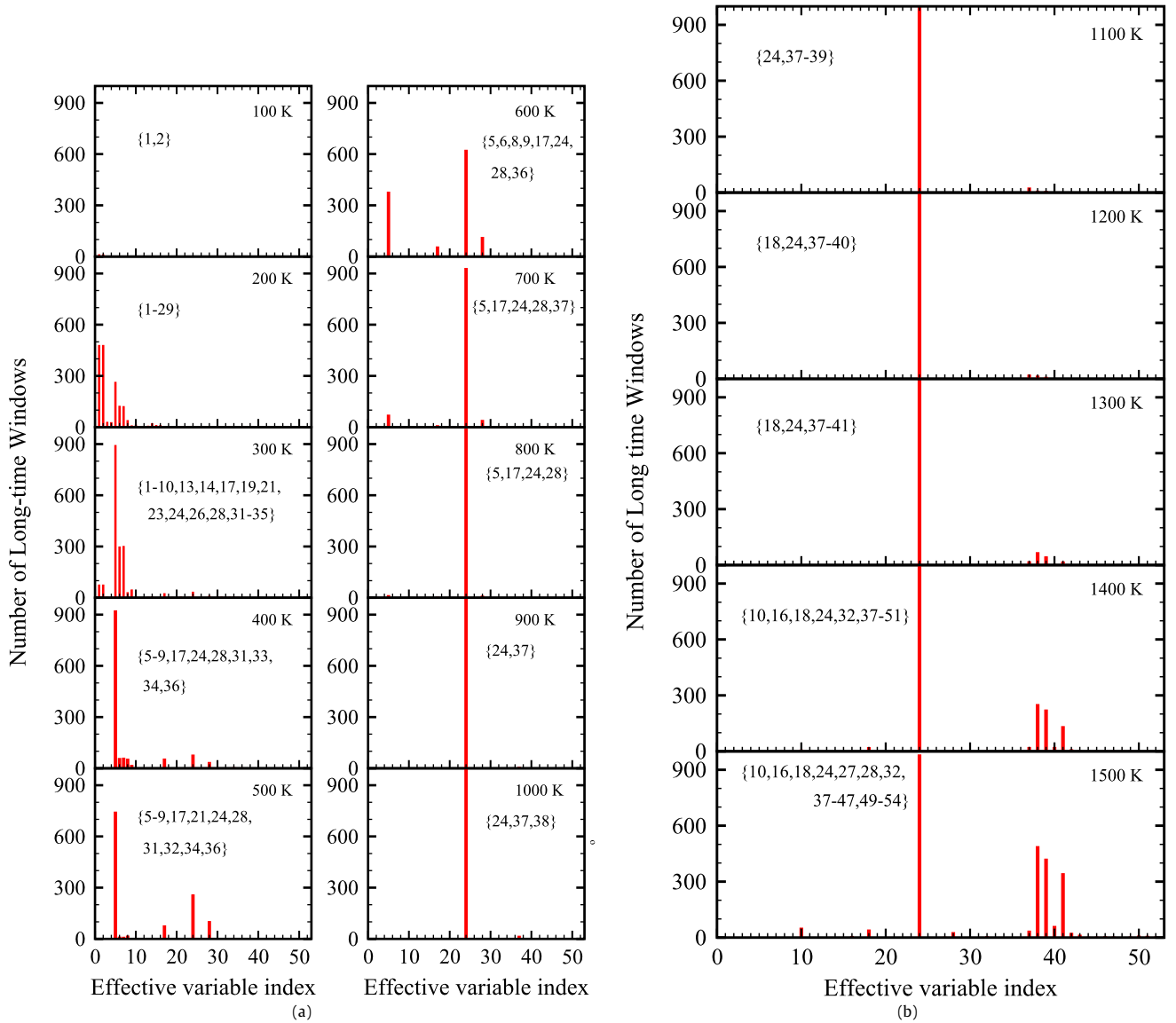


Fig. 9. Occurrence of effective variables in the total number of long-time windows for  $\text{Ag}_{14}$  for temperature range (a)  $T = 100$ – $1000$  K and (b)  $T = 1100$ – $1500$  K. In these figures, the integers in {...} are the effective variable index found at temperature  $T$ . Consult the look-up Table 2 for strongly-correlated CNLs corresponding to the effective variable index.

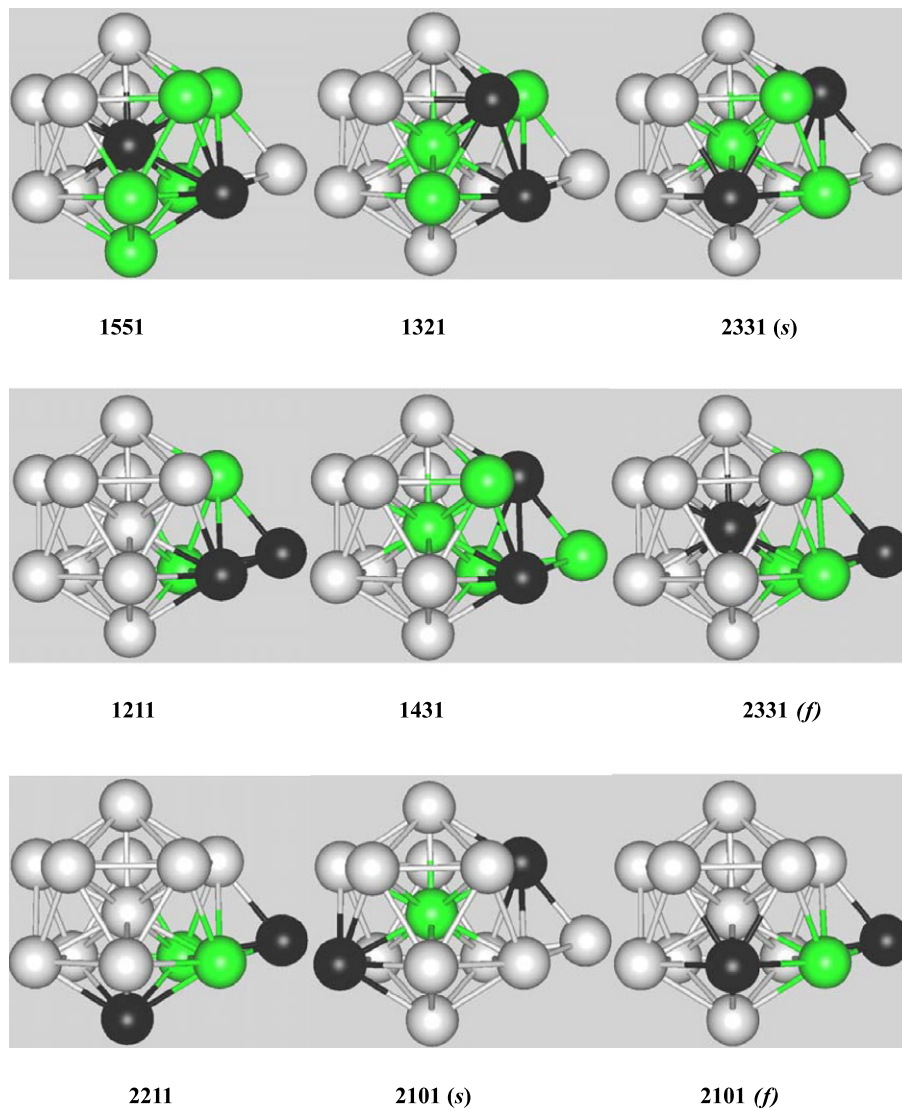


Fig. 10. Seven CNLs for the ground state configurations of  $\text{Ag}_{14}$ .

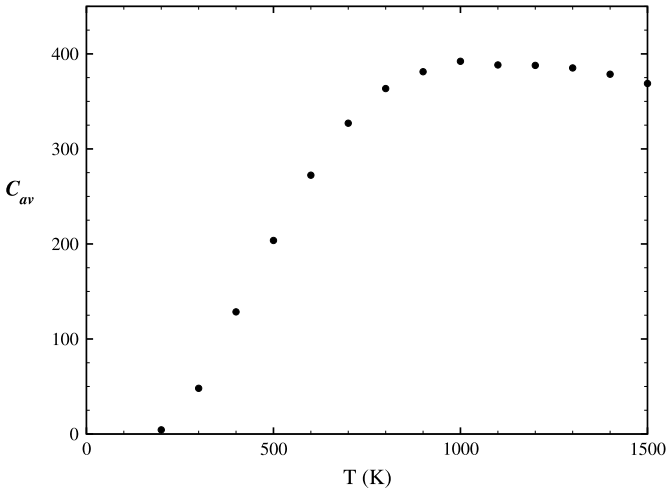
{1551} has one of its root-pair atoms which is the central atom of the icosahedron, it is reasonable to consider the CNL {1551} as maintaining a solid-like structure, and naturally interpret the disappearance of EVI-5 as indicating the cluster's tendency to melt. Based on this interpretation, the melting temperature possibly falls in the range  $800 \leq T_m \leq 900$  K. This melting temperature is less than  $T_m = 920$  K which we previously [2] deduced from the main maximum position of specific heat, but agrees with that we inferred from the analysis of the power spectral density (see Fig. 6 in Ref. [2]).

As pointed out above, EVI-24 appears more frequently with increasing temperature from  $T \geq 200$  K and is seen to replace EVI-5 for  $T \geq 900$  K. It would therefore be interesting to examine the correlation level matrix element {1321, 2331} in each of the long-time windows. To this end, we read off the value of this element in the correlation level matrices of all 999 long-time windows and denote their average value by  $C_{av}$ . Fig. 11 displays  $C_{av}$  plotted as a function of  $T$ . We find that there is a maximum at  $T \approx 960$  K.

### 3.2. Pure and bimetallic clusters: $\text{Cu}_{14}$ and $\text{Cu}_{13}\text{Au}_1$

We turn next to a comparative study of the time series analysis clustering for  $\text{Cu}_{14}$  and  $\text{Cu}_{13}\text{Au}_1$ . For  $\text{Cu}_{14}$ , the readers are again re-

ferred to Ref. [4] for its energy histogram (Fig. 3 in Ref. [4]) and the lowest energy and lower excited states (Fig. 2 in Ref. [4]), whereas for  $\text{Cu}_{13}\text{Au}_1$  its energy histogram is shown in Fig. 12 and its lowest energy and lower excited states are found in Ref. [1] (Fig. 5 in this reference). We note, first of all, that the lowest energy state of  $\text{Cu}_{14}$  persists up to  $T = 110$  K which is 40 K lower than that in  $\text{Cu}_{13}\text{Au}_1$ . The first excited state of  $\text{Cu}_{14}$  emerges at  $T = 120$  K, and there is no sign of the second excited state up till  $T = 500$  K. In contrast, the first excited state of the bimetallic cluster  $\text{Cu}_{13}\text{Au}_1$  occurs at  $150 < T \leq 160$  K, after which the second and third excited states appear and these two excited states are seen to grow all the way up to  $T = 500$  K. These differences in atomic configurations, which can be probed as functions of temperature, are reflected in subtle changes of effective variables. For  $\text{Cu}_{14}$ , one notices from Fig. 13(a) that EVI-1 which is {1101, 1201, 1311, 1541, 2201, 2211, 2321, 2431} and EVI-2 which is {1211, 1321, 1431, 1551, 2331} (see Table 3 to look up effective variables for  $\text{Cu}_{14}$ ) first climb up as the temperature increases from  $T = 100$  to 200 K after which they decline to about three times less in magnitude at  $T = 300$  K and vanish identically at  $T = 500$  K. The same trends are seen for  $\text{Cu}_{13}\text{Au}_1$  (Fig. 13(a)) except that EVI-1 which is {1101, 1201, 1311, 1541, 2201, 2211, 2321, 2431} and EVI-2 which is {1211, 1321, 1431, 1551, 2331} (see Table 4 to look

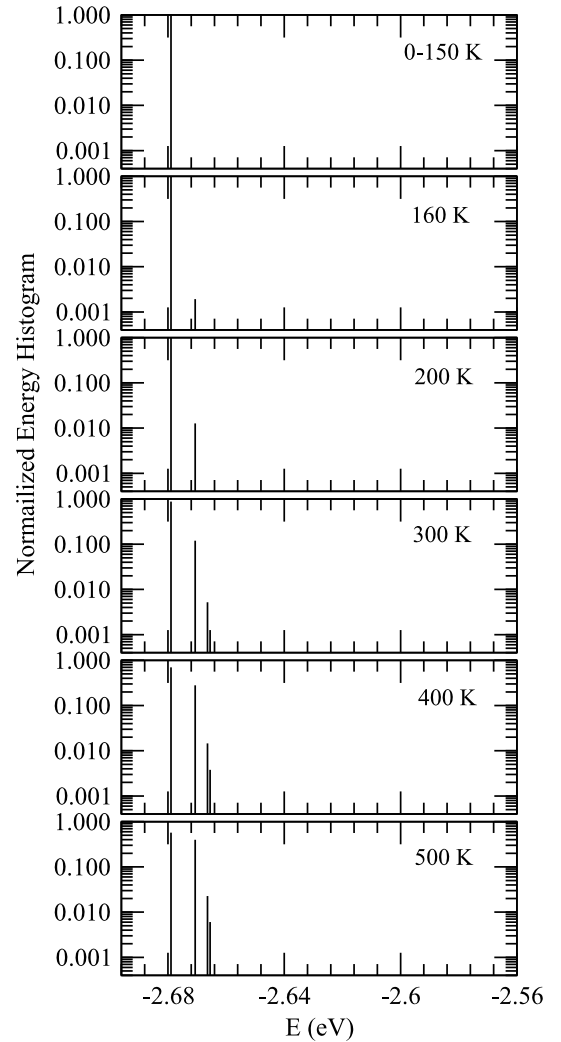


**Fig. 11.** Averaged value  $C_{av}$  (over all long-time windows) of correlation level elements {1321, 2331} that are read from  $C_{ij}$  vs.  $T$  (in Kelvin) for  $Ag_{14}$ .

up effective variables for  $Cu_{13}Au_1$ ) ascend more rapidly but also descend fast (for instance, at  $T = 300$  K they are already 7 times less in magnitude than those at  $T = 200$  K) and these effective variables disappear at  $T = 400$  K which contrasts to  $T = 500$  K in  $Cu_{14}$ . Another set of effective variables whose occurrence frequencies rise and fall with increasing temperature is EVI-11 and EVI-12 for  $Cu_{14}$ , and EVI-3 and EVI-4 for  $Cu_{13}Au_1$ . These effective variables peak between  $T = 300$  and  $400$  K for  $Cu_{14}$ , and at  $T = 300$  K for  $Cu_{13}Au_1$ . At  $T = 600$  K, they become vanishingly small both for  $Cu_{14}$  and  $Cu_{13}Au_1$ . Since the first and second excited states of  $Cu_{14}$  and with addition also the third for  $Cu_{13}Au_1$  emerge at these temperature ranges, the thermal variations of this second set of effective variables must be intimately related to the observed prepeaks of specific heats (at approximately  $400$  K for  $Cu_{14}$  and  $300$  K for  $Cu_{13}Au_1$ ). The temperature change of these effective variables agrees with the position of the prepeak maximum which is at higher temperature for  $Cu_{14}$  and lower for  $Cu_{13}Au_1$ .

Apart from the up and down characteristics of these two sets of effective variables, we see from Fig. 13(a)–(c) a conspicuous temperature response of the effective variable {1321, 1551, 2331} (EVI-3 for  $Cu_{14}$ , and EVI-5 for  $Cu_{13}Au_1$ , see Tables 3 and 4). These CNLs have the largest abundance values among the seven ground-state CNLs of both  $Cu_{14}$  and  $Cu_{13}Au_1$ . Just as in  $Ag_{14}$ , the occurrence frequency of this effective variable rises rapidly to its peak value between  $T = 400$  to  $500$  K in  $Cu_{14}$  and at  $T = 400$  K in  $Cu_{13}Au_1$ , and thereafter declines to identically zero at  $T = 1100$  and  $1300$  K, respectively. Following the argument used for  $Ag_{14}$ , the melting temperature is predicted to fall in the range  $1000 \leq T_m \leq 1100$  K for  $Cu_{14}$  and  $1200 \leq T_m \leq 1300$  K for  $Cu_{13}Au_1$ . For  $Cu_{14}$ , this melting temperature underestimates that inferred from the main peak position of  $C_V$  ( $T_m \approx 1257$  K, see Fig. 2 in Ref. [1]) by about  $150$  K. For  $Cu_{13}Au_1$ , the melting temperature predicted using the time series clustering method is closer to that deduced from the main peak position of  $C_V$  ( $T_m \approx 1357$  K, see Fig. 3 in Ref. [1]).

In closing this section, we would like to comment on the  $T_m$  obtained from the main maximum position of  $C_V$  and that calculated in the time series clustering method. In the former,  $T_m$  is an average temperature estimated by the drastic change of cluster's average potential energy over a finite temperature range, whereas in the latter it appeals to the abundances of CNLs whose values change temporally and rapidly with temperature. In the time series clustering method, the strongly-correlated CNLs are grouped together to become effective variables on the basis of the values of correlation level elements. For the presently considered clusters,



**Fig. 12.** Energy histogram for  $Cu_{13}Au_1$ .

the vanishment of an effective variable and its corresponding CNLs which we scrutinize over a specific range of temperature permits us to infer  $T_m$ , albeit approximately due to larger temperature intervals ( $100$  K) used in simulations.

#### 4. Conclusions

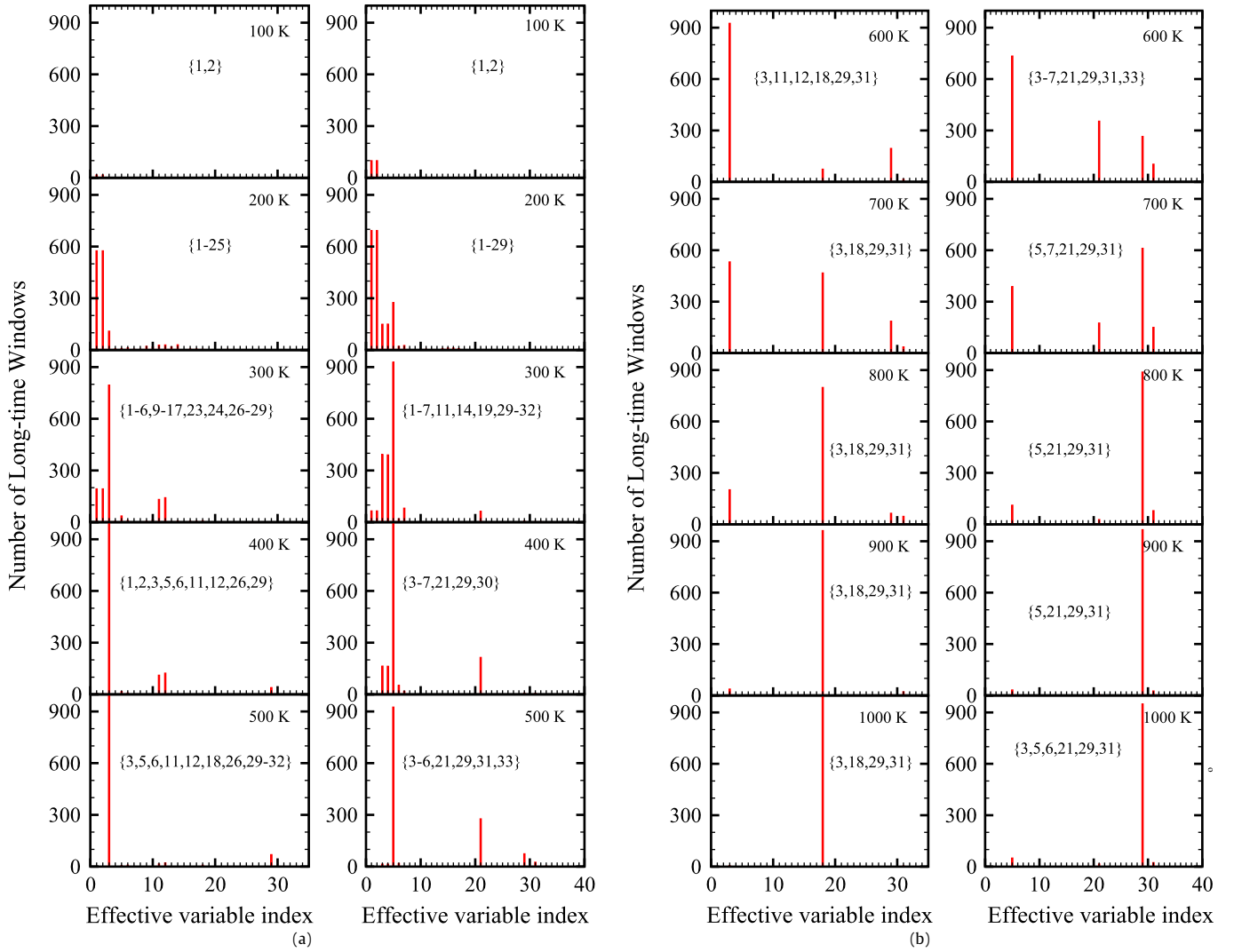
We have generated long-time series data sets for metallic clusters  $Cu_{13}$ ,  $Ag_{14}$ ,  $Cu_{14}$ , and  $Cu_{13}Au_1$  using an isothermal Brownian-type MD simulation. From the time evolution of atomic configurations recorded at different temperatures, the CNLs and their associated abundances were calculated. We performed statistical clustering on the change of CNL-abundance time series at a short-time scale, and then another statistical clustering based on the short-time correlation between CNL-abundance changes at a long-time scale. We found robust dynamical features in the short-time clustering using a short-time scale of  $\delta_{sw} = 10\Delta t$  ( $\Delta t = 0.001$  ps). However, we are at present unable to unambiguously select a long-time scale for the long-time clustering. Accordingly, we have chosen – by visual inspection of the short-time clustering density plots – to work with a long-time scale of  $2000\delta_{sw}$ . To test the reasonableness of this choice, we effected sliding window analysis on the simulation duration, advancing the long-time window at intervals  $1000\delta_{sw}$  to generate 999 overlapping long-time windows. This sliding window procedure was repeated for different temperatures. Using the long-time correlation matrix decomposition algorithm

**Table 3**Look-up table for the effective variables labeled by numeric integer index for Cu<sub>14</sub>.

Effective variable index	Number of CNL	CNLs							
1	8	1101	1201	1311	1541	2201	2211	2321	2431
2	5	1211	1321	1431	1551	2331			
3	3	1321	1551	2331					
4	2	1211	2321						
5	2	2211	2321						
6	5	1211	1431	1541	2211	2321			
7	6	1201	1311	1431	1541	2211	2441		
8	6	1321	1441	1551	1661	2331	2451		
9	2	1201	2211						
10	2	1541	2211						
11	2	1211	1431						
12	3	1541	2211	2321					
13	4	1101	2201	2321	2431				
14	3	1201	1311	2211					
15	2	1211	2331						
16	4	1201	1311	2321	2431				
17	2	2201	2211						
18	2	1321	2331						
19	4	1201	1311	1541	2211				
20	4	1321	1551	2101	2331				
21	3	1211	2101	2321					
22	6	1211	1431	1541	2101	2211	2321		
23	2	1201	1311						
24	2	2321	2431						
25	2	1211	2101						
26	4	1211	1541	2211	2321				
27	2	1551	2331						
28	3	1101	2201	2431					
29	3	1211	2211	2321					
30	4	1211	2101	2211	2321				
31	2	1211	2211						
32	3	1211	1431	2211					
33	2	1101	2101						
34	2	1311	2211						
35	2	1431	2331						

**Table 4**Look-up table for the effective variables labeled by numeric integer index for Cu<sub>13</sub>Au<sub>1</sub>.

Effective variable index	Number of CNL	CNLs							
1	8	1101	1201	1311	1541	2201	2211	2321	2431
2	5	1211	1321	1431	1551	2331			
3	2	1211	1431						
4	3	1541	2211	2321					
5	3	1321	1551	2331					
6	5	1211	1431	1541	2211	2321			
7	2	2211	2321						
8	6	1211	1431	1541	2101	2211	2321		
9	2	2101	2211						
10	2	1551	2331						
11	3	1201	1311	2211					
12	6	1201	1211	1421	2101	2211	2321		
13	4	1311	1321	1551	2331				
14	2	1201	2211						
15	4	1201	1311	2321	2431				
16	2	2201	2211						
17	2	1211	2321						
18	4	1321	1551	2101	2331				
19	4	1101	2201	2321	2431				
20	3	1101	2211	2321					
21	3	1211	2211	2321					
22	7	1101	1211	2101	2201	2211	2321	2431	
23	5	1211	1421	2101	2211	2321			
24	4	1201	1311	2211	2441				
25	2	1201	1311						
26	2	2321	2431						
27	4	1211	1541	2211	2321				
28	2	1541	2211						
29	2	1321	2331						
30	4	1211	2101	2211	2321				
31	2	1211	2211						
32	2	1541	2321						
33	3	1211	2101	2211					
34	2	1101	2101						
35	2	1311	2211						
36	3	1101	1201	2101					
37	2	1101	2201						
38	2	1201	2321						
39	2	1431	2331						



**Fig. 13.** Occurrence of effective variables in the total number of long-time windows for  $\text{Cu}_{14}$  (left column) and  $\text{Cu}_{13}\text{Au}_1$  (right column) for temperature ranges (a)  $T = 100\text{--}500\text{ K}$ , (b)  $T = 600\text{--}1000\text{ K}$  and (c)  $T = 1100\text{--}1500\text{ K}$ . In these figures, the integers in {...} are the effective variable index found at temperature  $T$ . Consult the look-up Table 3 for  $\text{Cu}_{14}$  and Table 4 for  $\text{Cu}_{13}\text{Au}_1$  for strongly-correlated CNLs corresponding to the effective variable index.

described in Section 2, we extracted effective variables (which are synchronized collections of CNLs) in each long-time window, and examined how frequently each effective variable occurring in the 999 long-time windows. It is observed that the effective variables show subtle thermal variations. In particular, we noticed a delicate change in the specific effective variable {1321, 1551, 2331} with temperature; this effective variable which is identified in all of metallic clusters considered here first increases with temperature, then declines, and finally vanishes identically at a high temperature. The disappearance of this effective variable is suggestive, for it indicates the cluster exhibiting a solid–liquid-like transition. We arrived at this interpretation based mainly on the fact that the CNL {1551} has one of the root-pair atoms resided at the central location of the icosahedron and structurally it bears most a solid phase. Apropos of cluster melting, we have made an attempt to deduce the melting temperature from our statistical time series analysis. The predicted melting temperature is always lower than the one inferred from the specific heat data. Specifically, the predicted melting temperature is reasonably good for  $\text{Ag}_{14}$  but falls short of for  $\text{Cu}_{14}$  and  $\text{Cu}_{13}\text{Au}_1$ . To understand this apparent discrepancy, we have experimented also with a smaller long-time window, and found the choice predicting higher melting temperatures. This implies that thorough and systematic investigations to

find an unambiguous way to decide what long-time scale to use for extracting effective variables at different temperatures are still in need.

As an epilogue, we note a relevant work of Kuntová et al. [33] who reported simulation studies of the melting process of several core–shell bimetallic clusters. Introducing the occupation probability  $P(T)$  (expressed in % at each temperature  $T$ ) of a cluster that takes on its lowest energy minimum over a simulation run, these authors describe the melting process as proceeding in two stages. The first stage involves defining a shell temperature  $T_m^{\text{layer}}$  above which  $P(T) < 80\%$ . In the second stage, another core temperature  $T_m^{\text{core}}$  is defined subject to the same prescription of requiring the core  $P_c(T) < 80\%$ . The temperature in the segment  $T_m^{\text{layer}} < T < T_m^{\text{core}}$  thus delineates the panorama where the external atomic shell layer melts whereas the internal core remains solid-like. It is attributed in their work that the melting process comes from the larger difference in energy between the lowest energy isomer and that isomer with higher energy. The temperature  $T$  in the range  $T_m^{\text{layer}} < T < T_m^{\text{core}}$  does not, however, agree with the  $T_m$  inferred from the principal peak position of caloric curve. We notice, on the other hand, that the main peaks of their caloric  $T_m$ s (Figs. 3 and 5 in Ref. [33]) comply very well with their  $P(T_m) = 0$  (Fig. 6 in Ref. [33] and re-calculated for the case



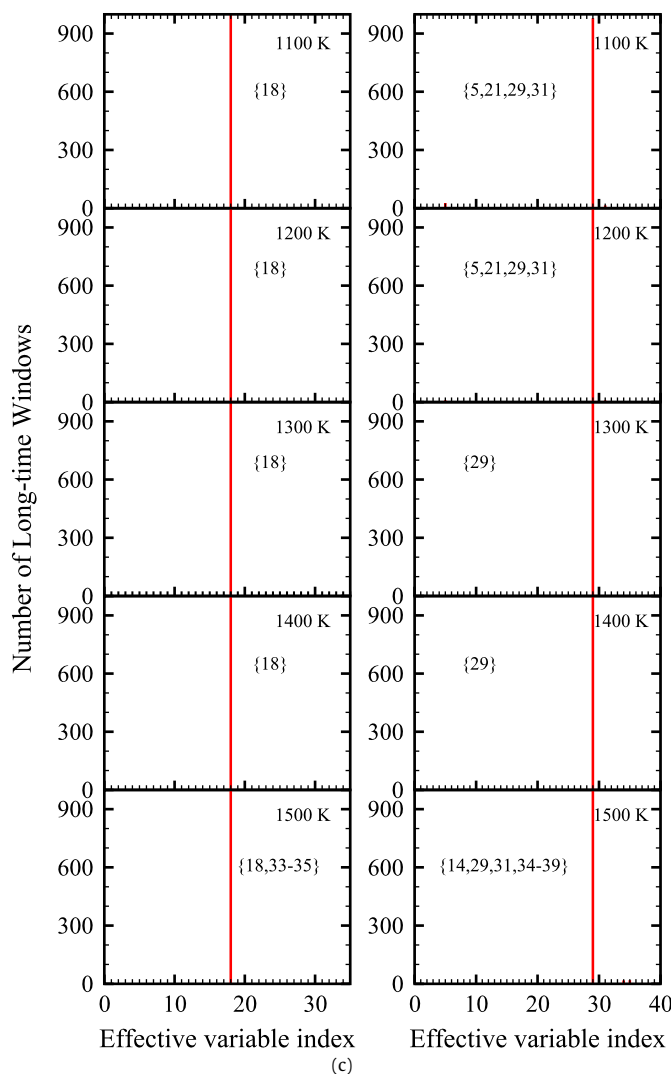
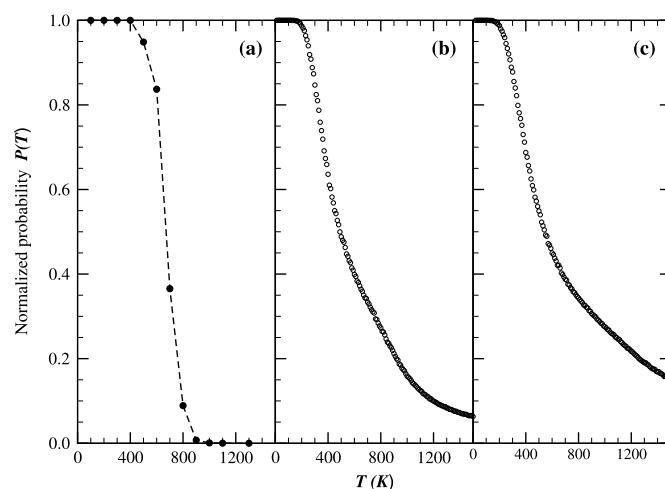


Fig. 13. Continued.

Fig. 14. Normalized occupation probability  $P(T)$  of finding the cluster in the lowest energy minimum for (a)  $\text{Ag}_{32}\text{Co}_{13}$ , (b)  $\text{Ag}_{14}$  and (c)  $\text{Cu}_{14}$ .

## References

- [1] Tsung-Wen Yen, S.K. Lai, N. Jakse, J.L. Bretonnet, *Phys. Rev. B* 75 (2007) 165420; S.K. Lai, W.D. Lin, K.L. Wu, W.H. Li, K.C. Lee, *J. Chem. Phys.* 121 (2004) 1487.
- [2] Tsung-Wen Yen, P.J. Hsu, S.K. Lai, e-J. Surf. Sci. Nanotech. 7 (2009) 149.
- [3] A. Rapallo, G. Rossi, R. Ferrando, A. Fortunelli, Benjamin C. Curley, Lesley D. Lloyd, G.M. Tarbuck, R.L. Johnston, *J. Chem. Phys.* 122 (2005) 194308.
- [4] P.J. Hsu, J.S. Lo, S.K. Lai, J.F. Wax, J.L. Bretonnet, *J. Chem. Phys.* 129 (2008) 194302. This reference summarizes some recent studies of clusters which are devoted to using velocity autocorrelation function.
- [5] V.V. Gafychuk, B.Yo. Datsko, J. Izmaylova, *Physica A* 341 (2004) 547.
- [6] A.M. Alonso, J.R. Berrendero, A. Hernández, A. Justel, *Comput. Statist. Data Anal.* 51 (2006) 762.
- [7] M.A. Juárez, M.F.J. Steel, CRISM Paper No. 06-14, Warwick University, 2006.
- [8] C. Allefeld, S. Bialonski, *Phys. Rev. E* 76 (2007) 066207.
- [9] L. Bauwens, J.V.K. Rombouts, *Econ. Rev.* 26 (2007) 365.
- [10] M. Corduas, D. Piccolo, *Comput. Statist. Data Anal.* 52 (2008) 1860.
- [11] S. Frühwirth-Schnatter, S. Kaufmann, *J. Business Econ. Statist.* 26 (2008) 78.
- [12] S.M. Focardi, F.J. Fabozzi, *Quant. Fin.* 4 (2004) 417.
- [13] M. Tumminello, T. Aste, T. Di Matteo, R.N. Mantegna, *Proc. Natl. Acad. Sci. USA* 102 (2005) 10421.
- [14] C. Rummel, G. Baier, M. Müller, *Europhys. Lett.* 80 (2007) 68004.
- [15] M. Tumminello, F. Lillo, R.N. Mantegna, *Europhys. Lett.* 78 (2007) 30006.
- [16] T. Heimo, G. Tibély, J. Saramäki, K. Kaski, J. Kertész, *Physica A* 387 (2008) 5930.
- [17] G. Innocenti, D. Materassi, *J. Phys. A: Math. Theor.* 41 (2008) 205101.
- [18] C. Rummel, *Phys. Rev. E* 77 (2008) 016708.
- [19] T.W. Liao, *Pattern Recognition* 38 (2005) 1857.
- [20] R.N. Mantegna, *Euro. Phys. J. B* 11 (1999) 193.
- [21] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, H.E. Stanley, *Phys. Rev. Lett.* 83 (1999) 1471.
- [22] S. Bialonski, K. Lehnertz, *Phys. Rev. E* 74 (2006) 051909.
- [23] U. Lee, S. Kim, K.-Y. Jung, *Phys. Rev. E* 73 (2006) 041920.
- [24] G. Baier, M. Müller, U. Stephani, H. Muhle, *Phys. Lett. A* 363 (2007) 290.
- [25] G.J. Ortega, R.G. Sola, J. Pastor, *Neurosci. Lett.* 447 (2008) 129.
- [26] A.K. Roopun, R.D. Traub, T. Baldeweg, M.O. Cunningham, R.G. Whittaker, A. Trevelyan, R. Duncan, A.J.C. Russell, M.A. Whittington, *Epilepsy & Behavior* 14 (2009) 39.
- [27] J. Basak, A. Sudarshan, D. Trivedi, M.S. Santhanam, *J. Machine Learning Res.* 5 (2004) 239.
- [28] M. Cogliati, P. Britos, R. García-Martínez, in: *Artificial Intelligence in Theory and Practice*, Springer, Boston, 2006, pp. 305–314.
- [29] S. Bivona, G. Bonanno, R. Burlon, D. Gurrera, C. Leone, *Physica A* 387 (2008) 5910.
- [30] Y.W. Goo, T.W. Lian, W.G. Ong, W.T. Choi, S.A. Cheong, *Financial atoms and molecules*, <http://arxiv.org/pdf/0903.2099>.
- [31] W. Wang, S.A. Cheong, A GPS time series clustering analysis of the 16 October 2007 Milford Sound Earthquake, presented at the 2010 Western Pacific Geophysics Meeting, 22–25 June, 2010, Taipei, Taiwan.
- [32] J.D. Honeycutt, H.C. Andersen, *J. Phys. Chem.* 91 (1987) 4950.
- [33] Z. Kuntová, G. Rossi, R. Ferrando, *Phys. Rev. B* 77 (2008) 205431.
- [34] P.J. Hsu, S.K. Lai, unpublished (2011).

## Acknowledgement

This work is supported by the National Science Council, Taiwan (NSC96-2112-M-008-018-MY3).