Eur. Phys. J. E (2011) **34**: 105
DOI 10.1140/epje/i2011-11105-9

THE EUROPEAN PHYSICAL JOURNAL E

Regular Article

Detecting hidden spatial and spatio-temporal structures in glasses and complex physical systems by multiresolution network clustering

P. Ronhovde[1], S. Chakrabarty[1], D. Hu[1], M. Sahu[1], K.K. Sahu[1,2], K.F. Kelton[1], N.A. Mauro[1], and Z. Nussinov[1,3,a]

[1] Department of Physics, Washington University in St. Louis, Campus Box 1105, 1 Brookings Drive, St. Louis, MO 63130, USA
[2] Metal Physics and Technology, ETH, 8093 Zurich, Switzerland
[3] Kavli Institute for Theoretical Physics, Santa Barbara, CA 93106, USA

Received 2 April 2011 and Received in final form 3 August 2011
Published online: 29 September 2011 –

Abstract. We elaborate on a general method that we recently introduced for characterizing the "natural" structures in complex physical systems via *multi-scale* network analysis. The method is based on "community detection" wherein interacting particles are partitioned into an "ideal gas" of optimally decoupled groups of particles. Specifically, we construct a set of network representations ("replicas") of the physical system based on interatomic potentials and apply a multiscale clustering ("multiresolution community detection") analysis using information-based correlations among the replicas. Replicas may i) be different representations of an identical static system, ii) embody dynamics by considering replicas to be time separated snapshots of the system (with a tunable time separation), or iii) encode general correlations when different replicas correspond to different representations of the entire history of the system as it evolves in space-time. Inputs for our method are the inter-particle potentials or experimentally measured two (or higher order) particle correlations. We apply our method to computer simulations of a binary Kob-Andersen Lennard-Jones system in a mixture ratio of $A_{80}B_{20}$, a ternary model system with components "A", "B", and "C" in ratios of $A_{88}B_7C_5$ (as in $Al_{88}Y_7Fe_5$), and to atomic coordinates in a $Zr_{80}Pt_{20}$ system as gleaned by reverse Monte Carlo analysis of experimentally determined structure factors. We identify the dominant structures (disjoint or overlapping) and general length scales by analyzing extrema of the information theory measures. We speculate on possible links between i) physical transitions or crossovers and ii) changes in structures found by this method as well as phase transitions associated with the computational complexity of the community detection problem. We also briefly consider continuum approaches and discuss rigidity and the shear penetration depth in amorphous systems; this latter length scale increases as the system becomes progressively rigid.

1 Introduction

This article constitutes a longer companion work to an earlier summary [1]. In this article, we elaborate on the details of our graph theoretical based method for the analysis of complex physical systems. We begin by briefly reviewing a special class of complex physical systems that is of great fundamental and technological importance — that of amorphous materials. From a practical standpoint, amorphous materials often have industrial processing and preparation advantages [2,3] relative to crystalline systems enabling, *e.g.*, greater solubility of pharmaceuticals [4] and many other advantages [2,5]. Below, we list several specific complex amorphous systems. i) Metallic glasses can be stronger than their respective crystalline structure and exhibit interesting electrical, chemical, and magnetic properties [3]. ii) Phosphate glasses are of great use in biomedical applications and chalcogenide glasses are of vital importance in optical recording media such as Blu Ray technology [6]. iii) Far more recent and exotic challenges involve incommensurate complex electronic structures found in systems such as the high-temperature superconductors [7,8]. Understanding the character of nontrivial structures in the above and many other systems is a problem of considerable interest in disparate arenas.

In perfect crystals, the natural system scales are evident by the regular ordering of the lattice. The fundamental unit cells of a crystal typically involve several atoms that are replicated in a simple pattern to span the entire system. There are no intermediate scale structures within the system from the atomic scale of the lattice up to the complete single crystal. Identifying the basic periodic unit cells is vital to the understanding of all crystalline solids. This simplistic structure enables an understanding of crystalline solids in great detail. Early on, the existence of spe-

[a] e-mail: zohar@wuphys.wustl.edu

cific unit cell structures was postulated to exist in crystals based on the sharp facets and other macroscopic properties of large crystals.

In more complex systems, new structures appear on additional intermediate scales between the atomic scale and the macro scale of the system. In recent years, scientific exploration has endeavored to understand a vast array of such complex materials that do not have a simple theoretical starting point.

As alluded to above, some of the best known complex materials are glasses. Liquids that are rapidly cooled ("supercooled") below their melting temperature avoid crystallization and instead become quenched into an amorphous state. On supercooling, liquids may veer towards *local* low energy structures (that cannot, on their own, be globally replicated to fill space without the inclusion of other structures) before being quenched into an amorphous state. Local low energy structures such as those formed by icosahedral packings [9] are indeed observed in metallic liquids [10,11]. Due to the lack of a simple crystalline reference, the structures of liquids and glasses are extremely hard to quantify beyond local scales.

Many glass theories rely on the hypothesis of natural structures in the glass [12–15]. However, actually finding such structures in a general way has been more elusive. *How, then, may we detect and best characterize the most notable structures in amorphous systems?* In the current work, we introduce a general framework to address this question. We illustrate the rudiments of our method by applying it to the analysis of glass formers.

The outline of the remainder of this article is as follows. We present some background information concerning the pursuit of characterizing structures in glasses and the basic features of our method in sect. 2. Details of the primary simulated systems are presented in sect. 3. The community detection and multi-scale community detections methods are explained in sect. 4. Our multiresolution method is applied to physical models in sect. 5. We summarize our findings in sect. 6. Details regarding the applied information measures are given in appendix A. Appendix B highlights the simple (yet often overlooked) fact that prepeaks in the structure factor do not constitute a necessary condition for medium range order.

In appendix C, we explain how our method allows for and treats *overlapping nodes* —nodes that are common to one or more "communities" within the community detection methods that we employ. In appendices D–L elaborate on additional test cases and various facets of our method. We conclude by discussing structures in space-time and their general properties. We discuss the detection of multi-scale structure in space-time in appendix M. In appendix N, we consider continuum approaches and discuss the divergence of a general length scale —the "shear penetration depth" (λ_{shear})— as a supercooled liquid becomes quenched into a rigid glass. At length scales smaller than λ_{shear}, the system may support shear and exhibit solid-like behavior. At larger length scales, the system exhibits liquid like behavior and shear stress is screened.

2 Background

2.1 Brief summary of numerous current approaches

Existing work in the pursuit of understanding the glass transition is vast. It spans many decades and is of importance across many fields of science and engineering. Glass formers exhibit several unique common features [16]. Glasses demonstrate short range order (SRO) and medium range order (MRO) structures, but no easily discernible static long range order exists. A striking property of glass formers (especially of the so-called "fragile" glass formers) is that their relaxation times (as measured, by, *e.g.*, their viscosity) can increase by many orders of magnitude over a relatively narrow temperature range. This dramatic slowing down is not associated with the usual distinguishing measures of conventional thermodynamic phase transitions. These systems have rich energy landscapes with an exceptionally high number of metastable states [17–23].

Given the broad appearance of glass-related states, different frameworks have been explored to work towards a "universal" characterization of the glass transition. Many theoretical approaches, *e.g.*, [15,14,16,24–28] have been developed over the years. The notable theory of random first order transitions (RFOT) investigates mosaics of local configurations [16,24]. RFOT is related [15] to theories of locally preferred structures [12–15,29]. Other theories seek similar measures of structure. Among many others, these approaches include spin glass type analysis [25], theories of topological defects and kinetic constraints [14,15, 30–32], and numerous ingenious approaches summarized in excellent reviews, *e.g.*, [20,33,34]. Formally, as demonstrated in [35], a growing static length scale is associated with the diverging relaxation times in supercooled liquids. Some works indeed found indications of increasing correlation lengths (static and those describing dynamic inhomogeneities) as the temperature was lowered [36–39]. Static correlation lengths were amongst other approaches, notably, examined in terms of i) "point-to-set" correlations [40,41] as well as ii) pattern repetition [42]. We briefly discuss these measures later on.

In metallic glasses, early work to ascertain local structures used a dense random packing model [43]. It was later established that such structures are better represented by an efficient cluster packing (ECP) model [44–46]. SRO features were thought to pivot on the existence of local icosahedral structures centered around solute atoms. Various idealized SRO configurations were presented in [47]. Schenk *et al.* and Kelton *et al.* [10,11] experimentally verified icosahedral short range order (ISRO) in undercooled liquids. Later work established the importance of ISRO in glasses [48–50].

Many structural characterizations are oriented toward static viewpoint of the system, but some dynamical features have also been examined. Analysis of "free volume" (unoccupied space between atoms) fluctuations [46] has been advanced. Several shear stress calculations explored dynamical processes in glass forming materials [22,51]. Dynamical heterogeneities (spatially non-uniform motion) in supercooled liquids have been investigated [52–56].

Viable characterizations of SRO and MRO structures were advanced for low [57] and high [50] solute concentrations, binary [50,57], and multicomponent systems [45].

Some methods of characterizing local structures include Voronoi tesselation [32,57,58], Honeycutt-Andersen indices [59], and bond orientational order parameters [60]. These specific measures center on local structures in the vicinity of an atom or a given link. As such, these specific measures (and correlations thereof) are restricted from detecting diverse complex long range structures.

Experimental means to directly measure MRO structures are given in [61,62]. Some potential MRO clusters were examined [23,50,57]. Several approaches to understand MRO use pattern matching to idealized MRO structures often constructed as agglomerations of perfectly ordered SRO features.

A very useful, experimentally driven, approach for ascertaining MRO looks for "prepeaks" in X-ray and neutron scattering data (that is, peaks in the structure factor $S(q)$ that appear for wave numbers q, corresponding to the inverse interatomic distances, which are lower in magnitude than that corresponding to the dominant $S(q)$ peak). We remark here that while this approach may capture general MRO structures, it is possible to have such structures without structure factor prepeaks (see appendix B).

2.2 Preliminaries concerning our method

Our unbiased structure characterization method extends *multiresolution* ideas [63] that have generally been reserved for network science applications (analyzing graphs in myriad social and biological networks) to complex materials. The key notion underlying the current work is that any complex physical system may be expressed as a network composed of nodes that code basic units of interest (*e.g.*, atoms, electrons, etc.). Weighted links may capture the strength of the interactions between the different nodes or experimentally determined correlations (*e.g.*, covariance or partial correlation contributions to the structure factor). After casting the system as a network, we then search for "communities" of nodes (*e.g.*, clusters of atoms) that are more tightly linked to each other (or —in the case of the use of correlation functions as link weights— are more strongly correlated with one another) than to nodes in other clusters [64].

Our multiresolution method employs the notion of "community detection", (*e.g.*, [65–76]), to quantitatively identify the "best" scale (or *scales*) for a complex physical system. Our approach *does not* rely on intuition or a knowledge of expected "important" features. Rather, it quantitatively estimates the best scale(s) through information-theory–based correlations, such as the variation of information (VI) [77] or normalized mutual information (NMI), among different solutions. (In appendix A, we review these information theory measures.) In essence, different copies of the community detection problem are given to different solvers ("replicas"). If many of these solvers strongly agree regarding certain features of the solution, then these aspects are more likely to be correct. Similarly, a large discrepancy may indicate large fluctuations on a particular scale (or scales). Extrema in NMI or VI among the results of these independent solvers then indicate the best scales for the network.

Multiple extrema can therefore indicate the existence of *multiple relevant length/time scales*. Although in most physical instances there is only a single dominant correlation length, there are many other cases in which more than one length/time scale is present [78].

One distinction between our work and some other established studies of local structures in glasses is that our analysis is not looking strictly at the positional structure. Rather, it evaluates structures in terms of the potential energies (*i.e.*, the internal binding energies of the clusters, see also [23]). Our method can encapsulate weights that represent general statistical (pair of higher order) stress (or other) correlation functions, relative atomic displacements, etc.

Our approach provides a perspective different from the "point-to-set" [35,79,80] and other methods [42]. The point-to-set method examines the overlap between configurations in a given volume (a "cavity") in an equilibrated system and compares those to configurations in the same cavity of the equilibrated system in which the boundary of the cavity was held fixed. Physically, it probes how probable it is to have a particular configuration within a disk or ball of a particular diameter given the boundary conditions. If many small clusters exist inside a sphere of fixed radius, then, a change in the boundary conditions will not significantly alter the bulk cluster distribution within the sphere. Conversely, if the sphere radius is smaller than the natural correlation length, then the number of configurations compatible with the boundary will be small and the overlap will be large. A different but perhaps related approach, that of ref. [42], examines the distribution of structures inside a given volume to identify the correlation length. The method examines whether the distribution of configurations inside the volume occurs with a random frequency (when the linear scale of the volume is *larger* than the correlation length) or not (when the linear scale of the volume is *smaller* than the correlation length).

Our method does not search for overlap at different scales for a multitude of configurations nor does it examine their frequency. Rather, the pertinent structures are naturally revealed by the information theory extrema between different copies of the entire system. The approach that we describe in this work does not require us to tabulate possible configurations and their occurrence frequencies nor does it require us to study the system in restricted volumes.

Furthermore, the basic structures that we find may be used as the natural units in a renormalization-group–type analysis or coarse-grained theories where clusters are replaced by single nodes and an effective energy can be written that entails interactions between the different clusters alone. As emphasized earlier, finding pertinent structures in amorphous systems such as structural glasses and numerous disordered systems is a non-trivial problem. In these systems, there is generally no symmetry or any other

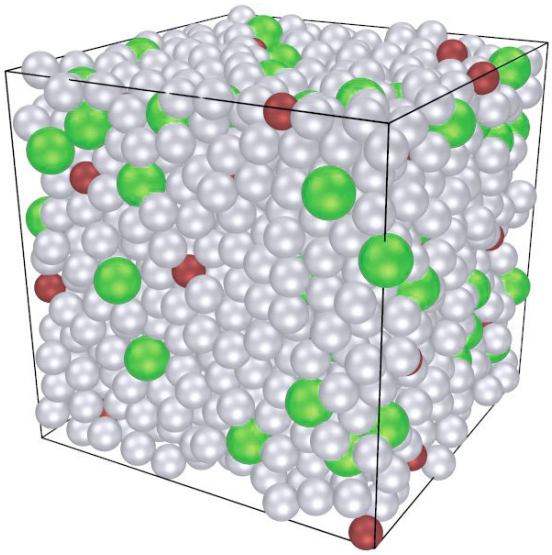

Fig. 1. (Color online) A depiction of our simulated model glass former with three components "A", "B", and "C" with mixture ratios of 88%, 7%, and 5%, respectively. The $N = 1600$ atoms are simulated via IMD [85] in a cube of approximately 31 Å in size with periodic boundary conditions. The identities of the atoms are C (red), A (silver), B (green) in order of increasing diameters.

obvious natural key that may determine how to optimally partition the system on different scales.

3 Simulations of model glasses

We examine a model glass former derived from a three-component AlYFe metallic glass [81] which we designate as "A", "B", and "C" in mixture ratios of 88%, 7%, and 5%, respectively. The presence of the different components B and C assists in the formation of a glassy state [82], since few pure compounds manifest a glassy state except under extreme preparation conditions. We additionally study the ubiquitous Kob-Andersen (KA) binary (80:20) Lennard-Jones type liquid [83] which lies in the glass-forming mixture region [84].

3.1 Ternary model glass former

In this section we discuss our study on the ternary system mentioned above. As depicted in fig. 1, we use classical molecular dynamics (MD) [85] to simulate the system dynamics. For this, we need accurate effective pair potentials that portray the pair-wise interactions between the atoms in the system. Our model potential energy function is [86]

$$\phi(r) = \left(\frac{a_0}{r}\right)^{a_1} + \frac{a_2}{r^{a_5}} \cos\left(a_3 r + a_4\right), \qquad (1)$$

where r is the distance between the centers of two atoms. This potential form incorporates a realistic weak long range interaction. Table 1 summarizes the parameter values a_i which depend on the specific types for a pair of

Table 1. Fit parameters for eq. (1) obtained from fitting configuration forces and energies to *ab initio* data [86,87]. The units of the parameters are such that, given r in Å, $\phi(r)$ is in eV. (That is, the parameters a_1, a_4 and a_5 are dimensionless, a_0 is in Å, a_2 is in eVÅ^{a_5} and a_3 is in Å^{-1}.) The same-species (*) data is replaced by a suggested potential derived from generalized pseudo-potential theory [88] (see also appendix C).

	a_0	a_1	a_2	a_3	a_4	a_5
AA	*	*	*	*	*	*
AB	1.92	17.4	6.09	3.05	−4.68	3.48
AC	2.38	8.96	−14.9	3.11	−3.88	4.38
BB	*	*	*	*	*	*
BC	1.88	8.00	−3.42	2.53	−1.25	3.00
CC	*	*	*	*	*	*

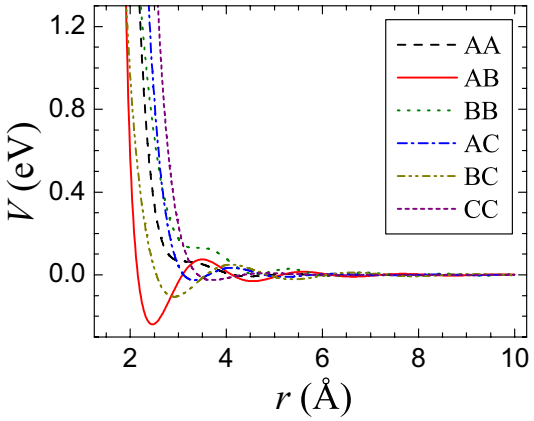

Fig. 2. The pair potentials for our three-component model glass former (see fig. 1). We indicate the atomic types by "A", "B", and "C" which are included with mixture ratios of 88%, 7%, and 5%, respectively. The units are given for a specific candidate atomic realization (AlYFe) discussed in the text. The same-species data uses a suggested potential derived from generalized pseudo-potential theory [88] (see also appendix C).

interacting atoms, and fig. 2 shows the respective potential plots.

The interaction parameters $\{a_i\}_{i=1}^{5}$ were determined [86] by fitting configuration forces and energies to *ab initio* data [87]. The same-species model interactions are finally replaced by that suggested by generalized pseudo-potential theory (GPT) [88]. As illustrated in fig. 1, we simulate $N = 1600$ atoms in a cubic system approximately 31 Å in size using periodic boundary conditions. This width is approximately twice the size of any suspected MRO structures.

The system is initialized at a temperature of $T = 1500$ K and allowed to equilibrate for a long time using a constant number of atoms (N), a constant volume (V), and a constant energy (E). That is, we work within the NVE ensemble. After allowing for system equilibration, we save s high temperature configurations separated by a fixed period of simulation time. (We fixed the value of s at 12 for all our simulations.) Prior to cooling, the length

scales in the system are changed by 1% to account for the increase in density as a result of cooling since we choose to cool the system in an NVT ensemble to control the temperature. This was done to keep a realistic density difference between the high and low temperature configurations. It should not have any important physical consequences pertaining to local structure. The system is then rapidly quenched to a temperature of $T = 300\,\mathrm{K}$, and it is allowed to equilibrate (in a mostly frozen state) in an NVE ensemble. We again save s separate low temperature configurations separated by a long period of simulation time.

3.2 Lennard-Jones glass

Here we discuss the application of our methods to simulations of the Kob-Anderson mixture. The pair potentials are given by

$$\phi_{\alpha\beta}(r) = 4\epsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{r}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r}\right)^{6}\right], \qquad (2)$$

where α or β designate one of two atomic types A and B. Specifically, in accord with KA we set the dimensionless units $\epsilon_{AA} = 1.0$, $\epsilon_{AB} = 1.5$, $\epsilon_{BB} = 0.5$, $\sigma_{AA} = 1.0$, $\sigma_{AB} = 0.80$, and $\sigma_{BB} = 0.88$.

As in the ternary glassy system above, we use MD [85] to simulate a LJ system of $N = 2000$ atoms. The system is initialized at a temperature of $T = 5$ (using energy units where the Boltzmann constant $k_B = 1$) and allowed to evolve for a long time. We save s high temperature configurations separated by 10000 time steps. The time step size is $\Delta t = 0.0069$ in LJ time units. Then, the system is rapidly quenched to a temperature of $T = 0.01$ which is well below the glass transition temperature of the KA-LJ system. The system is evolved in this mostly frozen state, and we save s low temperature configurations separated by 10000 steps of simulation time.

4 Multiresolution clustering on amorphous materials

Our idea is to apply, for the first time, multiresolution network analysis methods to ascertain *all* pertinent structures in complex amorphous materials. A key subcomponent of this analysis is the community detection method itself. We first explain these ideas in network analysis and their physical analogs.

4.1 Physical motivation

In an *ideal decomposition* of a large system into decoupled subsystems (communities), there is no interaction between different communities, and the system is effectively that of an ideal gas of disjoint communities. In the simplest setting in which the Hamiltonian would be block diagonal, the evolution of nodes (*e.g.*, atoms) in each community would be decoupled from all other nodes in other communities. We next consider a fundamental Newtonian many-particle setting. If the total force on a cluster is zero, $\boldsymbol{F}_{\mathrm{cluster}} = 0$ (*i.e.*, the cluster is decoupled from all others), then the particles in that cluster will drift (on average) at the center-of-mass velocity for that cluster. We note that in viscous systems, particles which experience a similar force will tend to move in unison [89].

In such instances, we may treat each community as a different particle in an ideal gas of non-interacting such particles. The general problem is to find (the time dependent) permutation that renders the pair interaction strengths and/or correlations between particles into the best possible block diagonal form (on the time scale chosen). Community detection emulates this for graphs. In the atomic realization that we discuss in this article, this emulates a partition of the system into optimally decoupled clusters such that the system may be viewed as an *"ideal gas" of decoupled communities*.

Slow cooling of a liquid enables crystallization which results in a first order or critical transition in the community detection problem. A similar transition materializes in many slowly cooled liquids as they crystallize to form ordered solids. By contrast, in an extremely rapid cooling of a liquid, the interactions between the particles are similar to those in spin-glass systems (*i.e.*, the particles are not nicely organized and consequently the inter-particle interactions harbor a large degree of randomness). It is notable that a spin-glass transition appears in the community detection problem for random graphs [90].

4.2 Community detection

In a graph such as that depicted in fig. 3, nodes correspond to abstracted fundamental elements of the system, and edges represent defined relationships between the nodes. Community detection describes the problem of finding strongly connected groups of nodes. Nodes in different clusters are more weakly connected than the nodes belonging to the same cluster. As alluded to earlier, we apply direct physical analogies where a node corresponds to a single atom. Edges, and their corresponding weights, are directly defined by the associated pair-wise interaction energy (or, in the absence of known interactions, may be set by measured inter-particle correlations). Specifically, we use the interatomic potential energies in eqs. (1) and (2). This potential model of the network edges is physically appealing in that finding the best partition for the network is akin to minimizing the cluster binding energies of the physical system. This is because our Hamiltonian favors solutions where the intra-cluster interaction energies outweigh inter-cluster energies. The clusters therefore behave as "solute particles" which interact weakly with the "solvent" and other solute particles. The problem is cast as identifying tightly bound clusters ("communities" or "solutes") against a background or "solvent" [91]. In the Potts model Hamiltonian that follows both attractive and repulsive interactions are present.

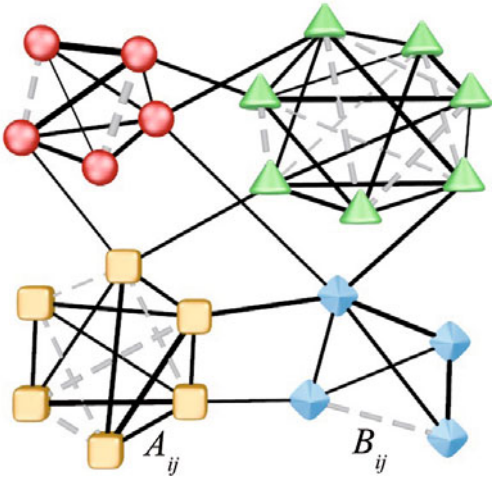

Fig. 3. An arbitrary weighted network with 4 natural communities (strongly connected) depicted as distinct node shapes. The goal in community detection is to identify any such strongly related clusters of nodes based on their defined edge relationships. Solid lines depict weighted links corresponding to complimentary or attractive relationships where $A_{ij} > 0$ and $B_{ij} = 0$ in eq. (3). Gray dashed lines depict missing, adversarial, or repulsive relationships where $A_{ij} = 0$ and $B_{ij} > 0$ in eq. (3). In both cases, the relative link weight is indicated by the respective line thicknesses. For presentation purposes, missing intercommunity edges are not depicted. In this paper, we directly relate the edges (attractive and missing/repulsive) to the interaction energy between pairs of atoms which implies that the natural groups would correspond to bound clusters of atoms.

Our Potts model Hamiltonian for community detection is [64]

$$\mathcal{H}(\{\sigma\}) = -\frac{1}{2}\sum_{i\neq j}\big(A_{ij} - \gamma B_{ij}\big)\delta(\sigma_i, \sigma_j). \qquad (3)$$

Globally minimizing this Hamiltonian corresponds to identifying strongly connected clusters of nodes. The elements of the matrices A_{ij} and B_{ij} are the edge weights and are defined as follows: an "attractive" weight has $A_{ij} > 1$ if nodes i and j are connected, $A_{ij} = 0$ if the nodes are not connected, a "repulsive" weight has $B_{ij} \geq 0$ if the nodes are not connected, and $B_{ij} = 0$ if nodes i and j are connected. We will set the weights A_{ij} and B_{ij} to represent the pair-wise potential energies (see sect. 5). These weights may similarly be set by pair-wise correlations. In principle, we can generalize the Hamiltonian to include n-body correlations or interactions.

We split the "attractive" (ferromagnetic) and "repulsive" (anti-ferromagnetic) contributions into two separate weighted matrices in order to insert the model weight γ that adjusts the energy trade-off between the two types of interactions. The parameter γ allows us to vary the target scale of the community solution. The spin states σ_i designate the community membership of each node i with a range $1 \leq \sigma_i \leq q$, where q is the number of communities. This number q may be variable [63,64] (such as in the multiresolution analysis that we perform in the current work) in order to find the optimal solutions or held fixed. Node i is a member of community k if $\sigma_i = k$. In this Hamiltonian, each spin σ_i interacts *only* with other spins in its *own* community.

Briefly, using eq. (3) our community detection algorithm rapidly moves nodes between communities based on the current lowest energy assignment until no more moves are possible (see also sect. 4.4). We then attempt p independent trials and select the lowest energy trial as the best division [63,64]. (The number of trials serves as our optimization parameter in this greedy algorithm. It effectively allows the algorithm to explore more possible configurations before selecting the best configuration for a given replica. The number of required trials in order to achieve a prescribed accuracy monitors the computational complexity (correlating with the number of local minima in which individual trials may get stuck). Somewhat better optimization could be obtained with a heat bath solution algorithm [90] at a cost of substantially increased computational effort.) This community detection algorithm partitions the network into communities by assigning a unique cluster membership for each node. Further details are provided in the appendices.

Local features in metallic glasses generally exhibit interconnecting short range structures [57]. In our community detection problem, this corresponds to allowing "overlapping" node memberships where atoms can be members of more than one local cluster. We incorporate this effect by assigning a node as a secondary member of every community for which it has a negative binding energy in terms our Potts model in eq. (3) (see appendix C).

4.3 Physical interpretation of community detection

Some previous analyses of local structure consider each cluster as "isolated" from the remainder of the bulk (simulation) liquid [9] or embedded within a liquid continuum approximation [91]. We pursue a direct analysis of the bulk (simulated) configuration where we adjust the cluster boundary (situated within the bulk material) based on the net Potts model interaction energy which is closely related to the simulated interaction energies. Specifically, a cluster membership for a given atom is explicitly defined when the atom has a negative Potts model (for the community detection problem) interaction energy with the entire cluster. In our tested configurations, this is related, but not explicitly equal, to the real binding energy with the other atoms in the cluster. The potential energy used for the edge weights is shifted by ϕ_0, and the Potts model energies for the "repulsive" edges (when a resulting edge weight is positive) are scaled by the model's resolution parameter γ (see sect. 4.5.1).

Larger clusters are affected more by surface interactions because the interactions with the distant atoms tend to zero, which corresponds to repulsive effects in the community detection problem (due to the ϕ_0 shift). A large collection of these edges makes it difficult for the atom to be assigned as a member of a particular cluster unless

the (near) surface interaction is strong. Coulomb-like interactions would likely display more limited surface-like effects in our analysis because of the slow decay rate over distance of the interaction.

An intuitive analogy for our analysis is that of a fragmentation model where the physical system is subjected to a large external impulsive force. That is, if we strike the system with an imaginary hammer, the system is more likely to fragment along surfaces with weak attraction along the walls of the boundaries. These lines intuitively and roughly correspond to the community boundaries due to how we define the network edges based on the potential values. This effect is seen in a simple example system that contains substantial defect sites [92]. Defects would likely fall along natural divisions after an impulse, and they appeared regularly along the boundaries of our community partitions.

Our analysis is independent of the type of structures that are being analyzed (structured, amorphous solid, and possibly even liquid systems). It is robust to noise in the model network, and it yields very accurate results with a simple and fast greedy algorithm [64]. Because edge assignments are based on relative node positions (through the interaction potential), our method should be robust with respect to translational or rotational motion of solid structures in the system (such as crystal nucleation).

4.4 Other physical community detection models

Minimizing the Hamiltonian of eq. (3) using the algorithm briefly explained above models a dynamic community detection process where we search for a local energy minimum that indicates a "good" community partition (in general). In particular, we have applied a potential energy (PE) model of network edges, and we could, in principle, apply other edge definitions to obtain other relevant configurations with the following caveat. When we minimize the Hamiltonian using the edge weights, a low energy state corresponds to a good partition, so for consistency with the community detection problem, any edge definitions should ideally be extremized at the most favorable configuration(s). The subtle distinctions between a few different natural edge weight models can result in different calculated clusters (beyond natural fluctuation caused by a high configurational entropy).

An inverted PE model fits the above criterion, and it would at first appear to be conceptually consistent with a model based on the relative pair-wise squared node displacement (SND) model. For both cases, small deflections about the minimum still indicate a good, even if not optimal, configuration corresponding to the intuitive notion of a bound cluster. However, the SND model would cause a perfect crystal to be identified as a single contiguous cluster. The PE would not have this effect since more distant nodes have a much lower potential energy. In effect, the PE model could identify the smaller scale units in a perfect crystal that the SND model could not isolate.

Another intuitive edge model for bound clusters is related to the attractive force that one may expect to exist between constituent elements of a physical cluster. Of course, the system forces still cause the formation of a crystalline ground under the appropriate conditions, but the forces dictate the system's physical *dynamics* over time where the PE model (more) directly indicates the relevant ground state(s). Forces with a minimum such as the LJ model in eq. (2) would be zero at the optimal ground state configuration (ideal crystal state), and the maximum force would be at a larger radius r than the PE minimum. While a force model of edge weights is also intuitive, it would correspond to (perhaps slightly) different clusters in a physical system compared to clusters derived from a PE model because the community detection Hamiltonian is extremized at different radii.

This discussion of the subtle differences between the edge models leads to the natural question of which is the "best" physical cluster model. Another perspective is that the different edge weight models simply answer different questions. We have selected the PE definition since it best corresponds to the ideal ground state of the system in our pair-wise interaction model. We then infer the best local clusters within configurations in local (frustrated) equilibrium (in the solid systems).

Our current analysis is not suitable for a hard-sphere system model because the energies are either zero or infinity. Similarly, repulsive-only potentials are not suitable for the current study because all atoms beyond a certain distance would be considered "attractive" in terms of the Potts model interaction resulting in a trivial cluster definition. Although for these cases, we could attempt other constructions for the community detection analysis such as that given in appendix L.

4.5 Multiresolution community detection

Multiresolution methods [63,93–96] extend the ideas of community detection to identify the "best" division(s) over a range of network scales ("resolutions"). We test s independent realizations of the system ("replicas") over all relevant network scales by specifying different values of γ in eq. (3). Networks with a weakly defined structure will result in more diverse solutions among the different replicas. Conversely, strongly defined structures will result in replica solutions that agree more strongly. When replicas represent time-separated configurations, a strong agreement among the replicas corresponds to consistent physical structures over time. The time separation between different snapshot can be tuned to find the corresponding pertinent structures for general time scales. The particular case of vanishing time separation corresponds to different representations of the same static system. See fig. 4. In the particular case of vanishing time separation, we permute, for identical nodes (atoms), the numbers labeling the nodes within the network; this leads to replicas corresponding to different initial starting points for the algorithm described in sect. 4.2. Another possibility, depicted in fig. 4, is that of replicas constituting different copies of the same system as it evolves in space-time (*i.e.*, a replica not corresponding to a given snapshot of the system at

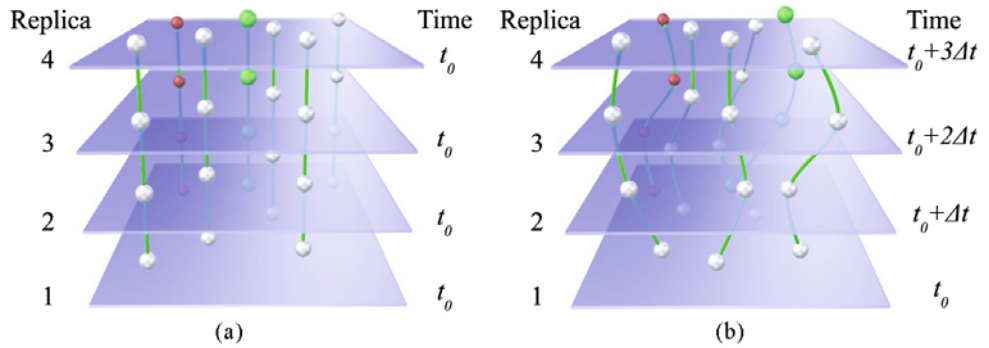

Fig. 4. Panel (a) depicts our replica construction for the physical system at a "static" time t_0 (no time separation between replicas, see appendix J). For presentation purposes, only a few nodes (atoms) are illustrated. Panel (b) depicts a similar set of replicas separated by a time Δt between successive replicas. In both panels (a) and (b), we generate a model network for each replica using the potential energy between the atoms as the respective edge weights for the network. Independently, within each replica, we then subsequently minimize eq. (3) at a given value of γ using the algorithm briefly described in sect. 4.2. Afterwards, we use the information measures in appendix A to evaluate how strongly the set of replicas agree on the best partition.

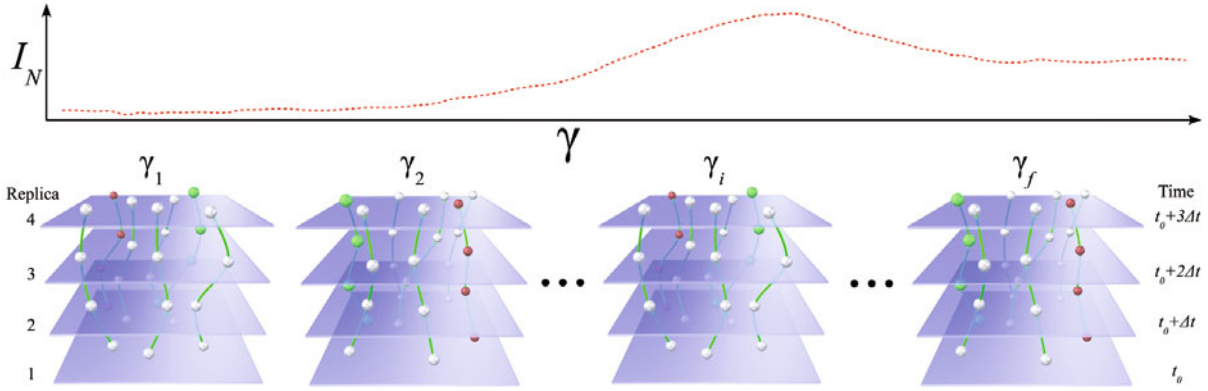

Fig. 5. A depiction of our multiresolution algorithm using the replicas of fig. 4(b) for a range of resolution parameters γ in eq. (3). We solve the set of replicas at each γ_i after which we utilize the information measures in appendix A, such as I_N in the schematic, to measure the level of agreement among the replicas for each tested resolution. The NMI I_N or VI V extrema (or plateaus in some instances) indicate preferred (or more "stable") resolutions.

a fixed time but rather monitoring the entire system as it evolves in space-time). This latter possibility will be alluded to in appendix M.

Within this framework, we identify the best resolution(s) by analyzing how well the independent replica solutions agree with each other in terms of information content [63]. See fig. 5 for a schematic involving time-separated replicas. Extremal correlated resolution(s) identify the best division(s) of the network. We apply the variation of information (VI) V metric and the normalized mutual information (NMI) I_N measure to evaluate the level of similarity among all pairs of the s replicas (see appendix A). Extrema correspond to locally stable solutions that remain locally unchanged for variations in the system scale. In the case of very stable system resolutions, local extrema can be replaced by plateaus in these information theory measures that indicate no change in the system solution over an extended range of resolution scales γ (as seen in the networks examined in [63]). We can further extract additional qualitative information about the "stability" of network partitions across a range of resolutions by examining the average number of clusters q [93,94,97], mutual information I [93], or the Shannon entropy H [63,97].

4.5.1 Physical interpretation of γ

The Potts model (resolution) parameter γ has a theoretical range of $0 < \gamma < \infty$. In practice, the upper range is generally not higher than $O(10^4)$ although the real upper and lower bounds depend strongly on the interaction energies. On a model level, γ simply scales the contribution of the repulsive edges for the community definition. The relevant values of γ are not absolute in the sense that they are readily compared among disparate problems, since the optimal quantities depend on the edge shift ϕ_0, the edge definition model, and other parameters of the system. In general, low values of γ (approaching zero) permit larger clusters until there is a single cluster (or completely disjoint subclusters). Conversely, "high" values of γ require smaller clusters, but the actual size of the minimal clusters depends strongly on the configuration and potentials.

In the LJ test case for example, the definitions of the clusters asymptotically approach weighted "clique" (maximally connected cluster) structures where every atom in a group is interacting strongly enough with the other members (a positive edge). That is, no matter how large γ is taken to be, the clusters do not break down into single atoms. Other community detection models will behave differently in this respect; but the effect is somewhat intuitive because, at the very least, the core of a well-defined physical cluster is a group of strongly bound atoms.

We can express the general partition function for a community partition with inter-community interactions which may correspond to the surface terms of clusters in Random First Order Transition theory (RFOT). Our parameter γ effectively plays the role of scaling the relation between surface and bulk terms in RFOT. A high value of γ corresponds to large surface effects while a small γ corresponds to dominant bulk effects.

Other physical system models (hard spheres) or forms of potentials (repulsive-only) would require additional consideration with corresponding modifications to our current clustering analysis. Varying γ has other effects for the current application such as adjusting the relative surface interaction dependence for the cluster determination (see sect. 4.3). We could also consider other physical (*i.e.*, not graph) distance scale parameters [98], where we can probe actual geometrical scales (not ones on an infinite-dimensional graph).

4.5.2 Phase diagram of the community detection problem

Apart from trying to find the best possible division of communities in a particular system, we have also looked at the ease at which such a solution can be found. This further indicates when, physically, the system might be analyzed in terms of nearly decoupled communities and when it cannot. It also enables us, by varying the parameters and temperature to map out when a solution is present or not, to determine whether the solutions found can be trusted and are physically relevant.

When solving the system of eq. (3) at non-zero temperature for a given network (*i.e.*, for an atomic configuration that is held fixed), entropic effects can, on their own, lead to a transition as the temperature is increased [90]. This transition corresponds to a spin glass type transition for random systems capturing a transition from easy to hard computational problems. These transitions may also be standard equilibrium transitions for sufficiently regular systems —critical for several regular graphs (*e.g.*, the square lattice viewed as a graph analyzed via a Potts model for $q \leq 4$ communities) or first order ($q > 4$ in the example above); for weak disorder, the latter transitions may be rounded off and further display signatures of Griffiths-type behavior [105]. We briefly elaborate on this idea below. As illustrated in the examples analyzed in this work, spatially increasing low temperature structures in a supercooled liquid as well as correlation lengths in systems such as the Ising spin systems can be ascertained via our

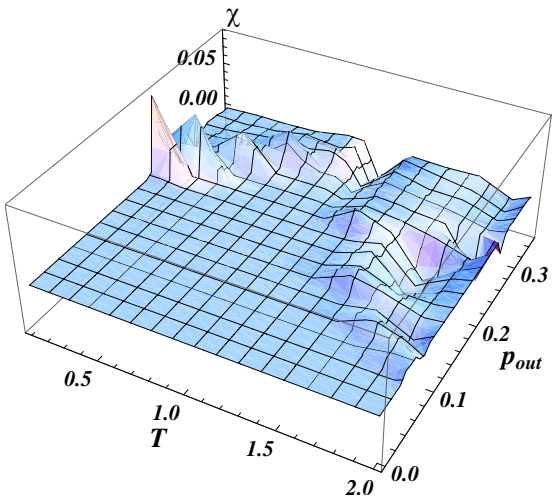

Fig. 6. A measure of the computational complexity χ of the community detection problem as a function of the temperature T and the density of inter-community links p_{out}.

method. In a general decomposition of an interacting system into optimally decoupled groups of particles ("communities"), the partition function is approximated as

$$Z \simeq \sum_{\{\Lambda\}} g(\Lambda) \prod_{c=1}^{q_\Lambda} z_c. \tag{4}$$

Here, z_c is the partition function as computed with the Hamiltonian of the entire system for the particles in community c, $g(\Lambda)$ is the frequency of obtaining a particular *set* of communities Λ in a decomposition of the entire system into optimally disjoint communities, and q_Λ is the total number of communities in the partition Λ. Similar considerations apply for quenched systems (where the free energies are weighted by $g(\Lambda)$). Thus, physical transitions or crossovers may, in this approximation, be related to transitions (or divergent scales) or crossovers in the communities found in Λ themselves and/or phase transitions associated with the computational complexity of the community detection problem (as further manifest via the distribution of partitions $g(\Lambda)$) [90]. A decomposition of the type of eq. (4) is exact for a (standard uniform) q state Potts model ($H = -\frac{1}{2}\sum_{i\neq j} A_{ij}\delta_{\sigma_i,\sigma_j}$ where $A_{ij} = 0,\ 1$) wherein the interaction energy between spins in different clusters (with each cluster/community given by uniform value of σ_i) is identically zero ($\delta_{\sigma_i,\sigma_j} = 0$).

In fig. 6, we plot a specific measure [63] of the computational complexity as a function of the density of energetically attractive inter-community links p_{out} and temperature T for a particular random graph. Within the solvable flat region (starting at low p_{out} and T) the system can be effectively decomposed into decoupled elements, *i.e.*, the partition function satisfies eq. (4) with a well-defined set of partitions $\{\Lambda\}$. This region is separated by ridges of high complexity in which the community detection problem becomes exceedingly hard from the unsolvable region in which no sensible community detection occurs (weak thermal fluctuations aid the system in avoiding metastable

states while large thermal fluctuations are detrimental). While the specific phase diagram boundaries were found with the Hamiltonian of our method (discussed next), the phase diagram changes little when other methods are used. In several simple cases the phase boundaries coincide with those of the known cases (*e.g.*, the phase transition of the Ising model on the square lattice when investigated as a network with links representing the strength of the spin exchange).

5 Multiresolution application to model glass formers

We assign edges between the nodes (atoms) with the respective weights based on the empirical pair potentials given by eqs. (1) and (2). Specifically, we calculate the potential energy ϕ_{ij} between each pair of nodes i and j in the system and then shift each value by a constant ϕ_0 to obtain $\phi'_{ij} = \phi_{ij} + \phi_0$ (assuming that $\phi_{ij} \to 0$ as the distance between particles i and j tends to infinity ($r \to \infty$)). The shift $\phi_0 > 0$ is necessary for the community detection algorithm to properly partition the network of atoms since it provides an objective definition of which interatomic spacings are preferable for a well-defined cluster and which are preferred to be excluded from a cluster.

In our particular application here, we calculate the average potential energy of the system and set $\phi_0 = -\phi_{\rm avg}$. For use in eq. (3), we define an edge with a weight $A_{ij} = -\phi'_{ij}$ between nodes i and j if $\phi'_{ij} < 0$, and we weight any missing links (or "repulsive edges") by $B_{ij} = \phi'_{ij}$ if $\phi'_{ij} \geq 0$. We then solve both model systems over a large range of γ using $s = 12$ replicas and $t = 10$ optimization trials per replica.

While $\phi_0 = -\phi_{\rm avg}$ is an intuitive shift that accomplishes the goal of an objective cluster definition here, it is not an appropriate shift for some problems. For example, using $\phi_0 = -\phi_{\rm avg}$ turns out to be problematic in some cases for lattice models. In a general setting, we examine a continuum of potential shifts ϕ_0 and monitor extrema in the information theory measures as a function of both γ in eq. (3) and ϕ_0.

In addition to the systems tested below, we applied the algorithm to various test cases including square, triangular, and cubic lattice structures. The algorithm is able to correctly identify the natural leading order scales (plaquettes and composites of plaquettes as "cascades" in the information theory correlations). Further testing involved two-dimensional defects (dislocations, interstitials, etc.) and domain walls in a lattice. Defects in triangular lattices occurred most frequently near cluster boundaries.

We also tested static configurations for the ternary model glass system where each replica is a model of the same configuration. There we detected structures in both low and high temperatures where the high temperature "structures" are more fragile (that is, harder to solve in the clustering problem). This corresponds to identifying relevant transient features in a dense liquid.

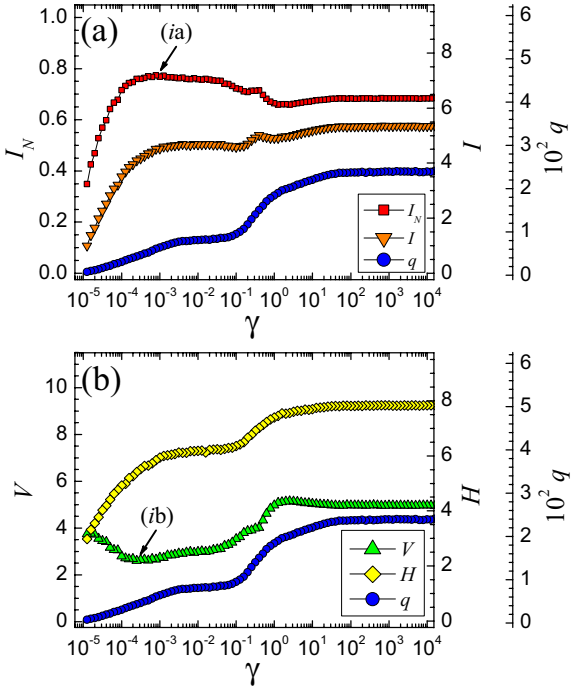

Fig. 7. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). The ternary model system contains 1600 atoms in a mixture of 88% type A, 7% of type B, and 5% of type C with a simulation temperature of $T = 300\,\mathrm{K}$ which is well *below* the glass transition for this system. This system shows a strongly correlated set of replica partitions as evidenced by the information extrema at (i) in both panels. A set of sample clusters for the best resolution at $\gamma \simeq 0.001$ is depicted in fig. 11.

5.1 Ternary model glass results

In figs. 7 and 8, panels (a) and (b) show the information theory based correlations (averaged over all replica pairs as in [63]) over a range of network resolutions. The lower temperature system at $T = 300\,\mathrm{K}$ in fig. 7(a) shows a peak NMI at (ia) with a corresponding VI minimum at (ib). Figure 9 depicts an example of the full system *partition*. Figure 10 shows some sample clusters within the simulation bounding box at resolution parameter value of $\gamma_{\rm best} \simeq 0.001$, where we include overlapping node memberships (the replicas correlations are calculated on partitions), and fig. 11 depicts additional samples of the best clusters. The corresponding high temperature ($T = 1500\,\mathrm{K}$) solutions have a much lower NMI at $\gamma_{\rm best} \simeq 0.001$ indicating significantly worse agreement among replicas. That is, one would expect that the high temperature system $T = 1500\,\mathrm{K}$ is in a liquid state, so any observed features are not dynamically stable across all replicas (snapshots of the system over time). At $T = 300\,\mathrm{K}$, the best structures have consistent cluster sizes that are exclusively MRO.

The plateau regions for $\gamma > 10$ are similar to the LJ plot in fig. 12, but in this system the NMI plateau is lower. In the high temperature case in fig. 8, there are additional "almost-plateaus" for the range $0.001 \lesssim \gamma \lesssim 0.1$. These

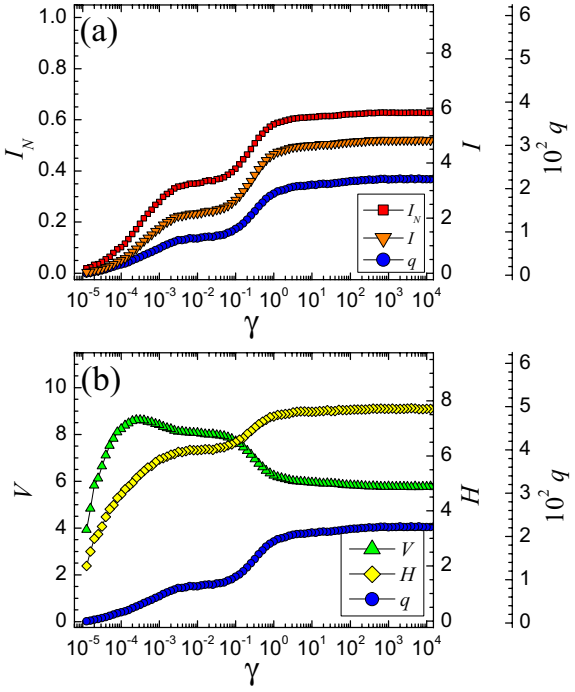

Fig. 8. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). The ternary model system contains 1600 atoms in a mixture of 88% type A, 7% of type B, and 5% of type C with a simulation temperature of $T = 1500\,\mathrm{K}$ which is well *above* the glass transition for this system. At this temperature, there is no resolution where the replicas are strongly correlated. See fig. 7 for the corresponding low-temperature case where the replicas are much more highly correlated at $\gamma \simeq 0.001$.

plateaus represent a region of structural transition, but we are not concerned with them because the replica correlations are very low.

5.2 Binary Lennard-Jones glass results

In figs. 12 and 13, panels (a) and (b) show the data for the replica information correlations over a range of network resolutions. The lower temperature system at a temperature of $T = 5$ (in units of $k_B = 1$) in fig. 7(a) shows a plateau in NMI at (ia) with a corresponding VI plateau at (ib) which are the local extrema ($V = 0$ is a trivial solution with only one cluster in this problem). Figure 14 depicts a sample of the best clusters, including overlapping node memberships, at resolution (i) for $\gamma_{\mathrm{best}} \simeq 10^4$. The corresponding higher temperature solutions at $\gamma_{\mathrm{best}} \simeq 10^4$ (see figs. 13 and 15) have a lower NMI (indicating a weak agreement among replicas). The dependence number of replicas (see fig. 15) required to achieve high accuracy underscores the faint agreement between contending solutions and the high temperature complexity of the problem. Our identified structures for this LJ model system are consistent in terms of the cluster sizes and are almost exclusively SRO

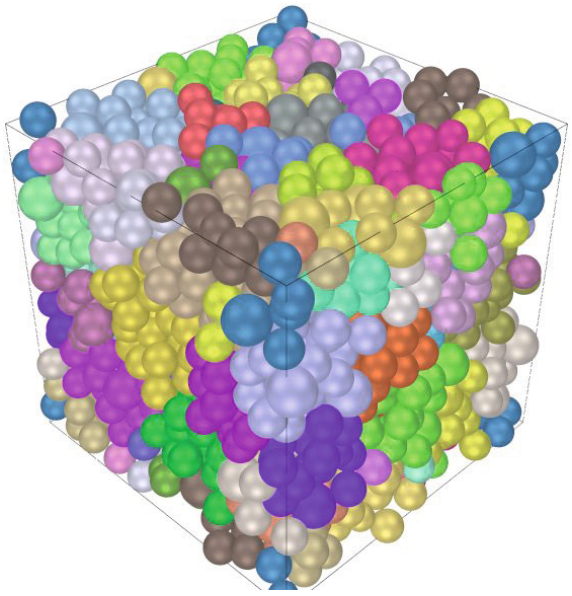

Fig. 9. (Color online) A depiction of the full *partitioned* system where unique cluster memberships are depicted as distinct colors (best viewed in color). The atomic identities are B, A, C in order of increasing diameters. Overlapping nodes (multiple memberships per node) are added to these communities to determine the best interlocking system clusters.

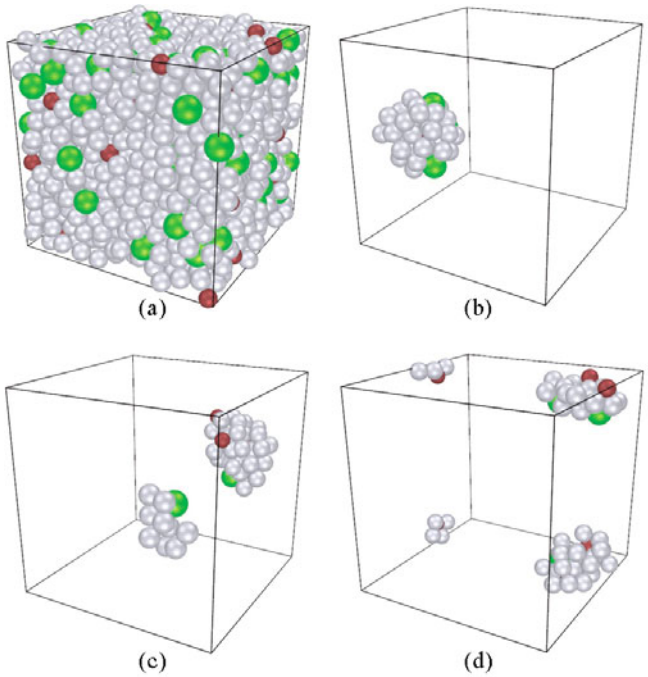

Fig. 10. Panel (a) is the full system cube, and panels (b)–(d) show three sample clusters (one distinct cluster each using periodic boundary conditions) within the simulation box. Note that the algorithm can identify structures beyond immediate short-range neighbors.

configurations with simple adjunct-type atoms extending into the low end of MRO size structures.

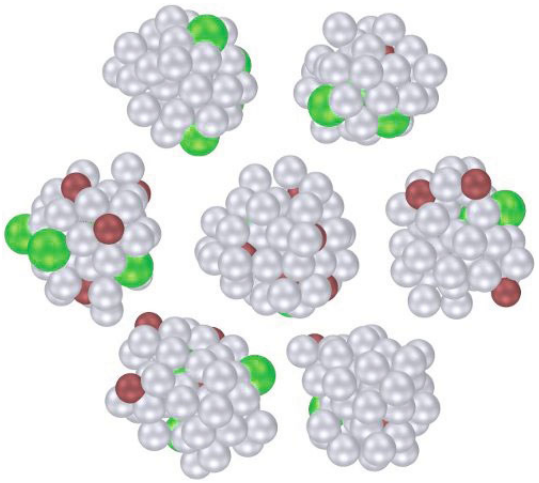

Fig. 11. (Color online) A depiction of some of the best clusters of the low temperature ($T = 300\,\mathrm{K}$) ternary system at the peak replica correlation at feature (i) in fig. 7. These clusters include overlapping node membership assignments where each node is required to have an overall negative binding energy to the other nodes in the cluster. The atomic identities are C (red), A (silver), B (green) in order of increasing diameters.

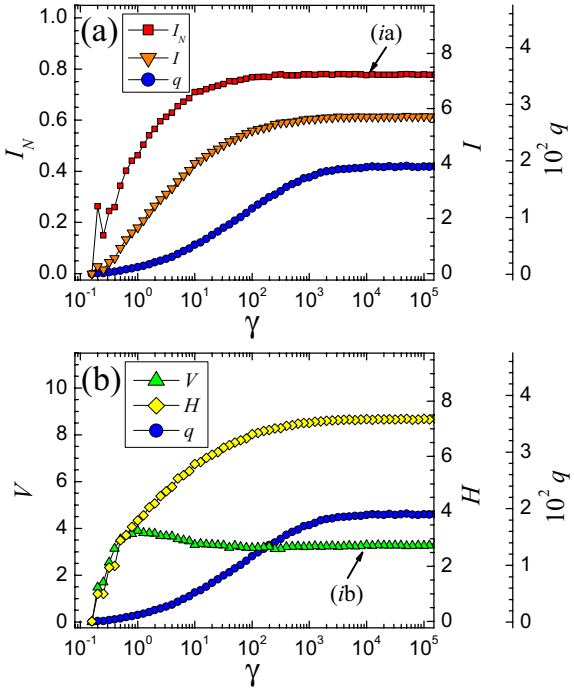

Fig. 12. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). The LJ system contains 2000 atoms in a mixture of 80% type A and 20% type B (Kob-Andersen binary LJ system [83]) with a simulation temperature of $T = 0.01$ (energy units) which is well *below* the glass transition of $T_c \simeq 0.5$ for this system. This system shows a somewhat strongly correlated set of replica partitions as evidenced by the information extrema at (ia,b) in panels (a) and (b). A set of sample clusters for the best resolution at $\gamma = 10^4$ is depicted in fig. 14.

6 Conclusions

Our method constitutes a new and very general approach to determine the natural multi-scale structures of complex physical systems. We do not bias the expected configurations in any way. The required input is that of inter-particle interactions (or measured correlations as further detailed in appendix L). Information theory extrema (including plateaus) between contending solutions give the different pertinent structures on all important length scales (lattice scales and correlation lengths) of the system in an unbiased unified way.

To illustrate the feasibility of this approach, we focused in this work on structural glasses. The detection of structure in structural glasses is a heavily investigated hard problem. By the use of our method, we identified consistent SRO or MRO structures at temperatures below the glass transition in two different model glass formers. Even though this does not imply general applicability to all glass formers, our approach introduces a new technique for determining the most natural structural features in some amorphous systems, and we leave the task of verifying its general applicability to future investigations.

Our analysis evaluates structures in terms the potential energies (*i.e.*, the internal binding energies of the clusters). This approach differs from some other methods of structural analysis that look strictly at the relative atomic positions. We briefly comment on the relation between our work and that of simpler mode analysis of the interactions —the latter have been indeed been used to detect a link between spatial structure and local dynamics [54, 99]. Local heterogeneous dynamics was shown to be correlated with topological defects [100] and thus (as a consequence of [54,99]) with the system modes. (Indeed, a recent work [101] directly reaffirms such a connection.) Similarly, some old variants of graph partitioning methods such as direct spectral clustering [102,103] as well as community detection employ a normal mode analysis [72]. Indeed, a method based on particle dynamics in high dimensions enabled multi-scale community detection [73]. Other different yet conceptually related approaches also include oscillator synchronization analogies [74,104]. In the above approaches, physical analogies were made. It is thus natural to suspect that the approach may be inverted and that community detection will link spatial structure with dynamics in a broad class of physical systems where the physics based analysis of the interacting multi-particle system is hampered by the shear complexity of the problem. The current work indeed elucidates this link in detail and introduces a method for the direct detection of general spatio-temporal structures in rather general physical systems. Notably, via the use of information theory correlations and extrema as a function of the Hamiltonian parameters therein (γ of our defining Hamiltonian of eq. (3) and ϕ_0 of sect. 5), we are able to identify in an unbiased way *all of the natural scales* of the system. In more rudimentary approaches this needs to be introduced by

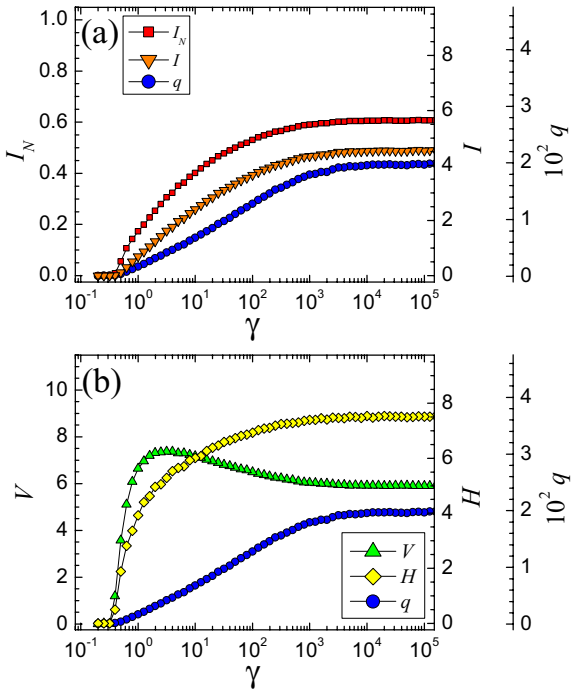

Fig. 13. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). The LJ system contains 2000 atoms in a mixture of 80% type A and 20% type B (Kob-Andersen binary LJ system [83]) with a simulation temperature of $T = 5$ (energy units) which is well *above* the glass transition of $T_c \simeq 0.5$ for this system. At this temperature, the replicas are significantly less correlated than the corresponding low temperature case in fig. 12.

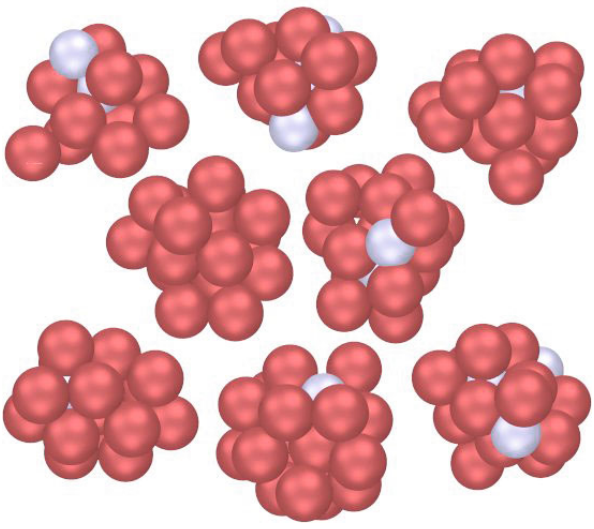

Fig. 14. (Color online) Several of the best clusters for the peak replica correlation at feature (i) in fig. 12. These clusters include overlapping node membership assignments where each node is required to have a overall negative binding energy to the other nodes in the cluster. The atomic identities are B (silver) and A (red) in order of increasing diameters.

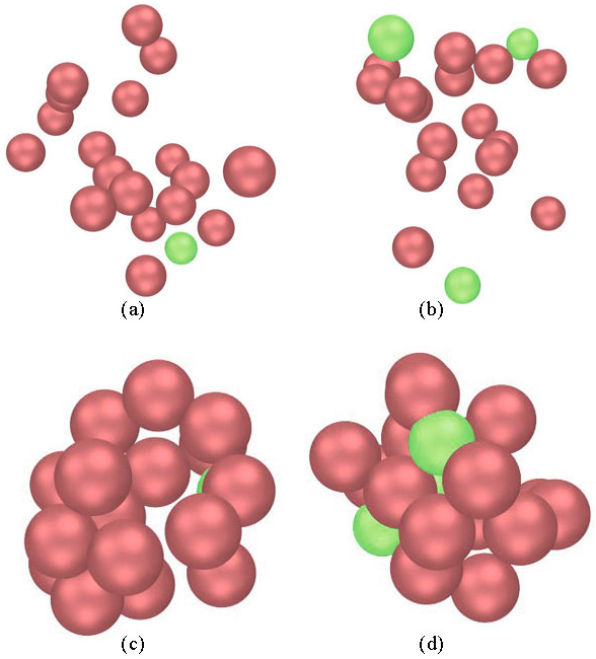

Fig. 15. (Color online) A sample depiction of dispersed clusters for the LJ system eq. (2) at a temperature of $T = 5$ (in units where $k_B = 1$). The shown clusters correspond to the multiresolution plot in fig. 13 at a value of the resolution parameter of $\gamma = 10^4$. These clusters are a sample of high temperature counterparts to the low temperature clusters in figs. 12 and 14. Panels (a) and (b) show a more typical example of dispersed clusters at a number of replicas $s = 10$. In some cases, the identified high temperature clusters can be more compact, but not densely packed. Panels (c) and (d) provide sample solutions for $s = 20$ replicas. An increasing replica (s) number (and generally also trial (p) number, see text) required to achieve better solutions is indicative of a greater computational complexity of the system. Physically, a larger time is required for the system to realize better clusters under ideal conditions (in the absence of quenching and any relaxation time constraints). *The sparsity of the identified clusters in this high temperature system is generally consistent across all clusters in the network solution.* (This lies in contrast to the more compact and more strongly correlated structures found at low temperatures.) The atomic identities are B (silver) and A (red) in order of increasing diameters.

hand (*e.g.*, a cutoff on mode occupancy in spectral based approaches) and it is not obvious how to determine all pertinent structures of a complex physical system.

Our approach identifies MRO as the dominant feature of our ternary model glass former with no strongly defined SRO. In contrast, the LJ system shows a largely SRO structure with adjunct atoms that create near-MRO structures. Augmenting the changes in structure that we find by analyzing the atomic system at different temperatures and minimizing the energy function to determine the optimal division into clusters, there are also *entropic effects*. The distribution of optimal partitions becomes wider and less pronounced due to these effects as the temperature increases. This is a validation of our arguments in sect. 4.5.2.

On a lattice, plateaus in information theory correlation steps correspond to a cascade of structures starting from the smallest dyads of nodes, to basic plaquette structures (square, triangle, etc.), and growing ever larger (two joined plaquettes etc.). In Ising spin systems at different temperatures on a square lattice, the domains of "+" and "−" spins are separated from one another by domain walls. The information theory plateaus correspond similarly to the cascade of small plaquette structures found on the lattice itself (*i.e.*, the single plaquette, two joined plaquettes, etc.) up to a cutoff scale set by the domain wall. This is sensible since no clear structure is found beyond the domain length scale. The largest fluctuations occur at the boundaries between different domains. These domain walls are directly attained by the extrema (those corresponding to the *maximum* in VI). Physically, they correspond to the scales at which the largest fluctuations occur where the large fluctuations lead to poor information theory correlations between the different replicas.

As we detail in appendix N, a general rigid amorphous solid supports shear and a divergent "shear penetration length". This length scale monotonically increases as a liquid is cooled to form a rigid glass. Within our graph theoretic method, we may employ the shear stress correlations as to represent graph weights to analyze such behavior.

We are indebted to M. Widom and M. Mihalkovič for help with effective atomic potentials and to ongoing work [106]. This work was supported in part by the LDRD DR on the physics of algorithms at LANL and the Center for Materials Innovation at Washington University in St. Louis. ZN thanks N. Goldenfeld, Z. Rotman, G. Tarjus, and P. G. Wolynes for helpful comments. He also wishes to thank the KITP and the Lorentz Center for hosting stimulating workshops. ZN gratefully acknowledges the National Science Foundation (NSF-DMR-DMR-1106293). This research was further supported in part by the National Science Foundation under Grant No. NSF PHY05-51164 (KITP).

Appendix A. Information theory measures

We utilize the variation of information [77] and normalized mutual information to measure the strength of the correlations among the independent replicas in our multiresolution algorithm. The mutual information $I(A,B)$ between community divisions A and B is

$$I(A,B) = \sum_{i=1}^{q_A}\sum_{j=1}^{q_B} \frac{n_{ij}}{N} \log\left(\frac{n_{ij}N}{n_i n_j}\right). \tag{A.1}$$

q_A and q_B are the number of communities in divisions A and B, n_i and n_j are the number of nodes in communities i and j, and n_{ij} is a "confusion matrix" that identifies the number of nodes in community i of partition A that are found in community j of partition B. For a single community division A, we can determine the self-mutual information by $H(A) = I(A,A)$ which is identically equal to the Shannon entropy. We use base 2 logarithms.

We calculate the variation of information $V(A,B)$ by

$$V(A,B) = H(A) + H(B) - 2I(A,B). \tag{A.2}$$

The range for VI is $0 \leq V(A,B) \leq \log N$. The normalized mutual information $I_N(A,B)$ is

$$I_N(A,B) = \frac{2I(A,B)}{H(A)+H(B)}\,. \tag{A.3}$$

The range for NMI is $0 \leq I_N(A,B) \leq 1$. The NMI and VI afford slightly different perspectives on the replica correlations in the multiresolution analysis above.

These information measures are based on the cluster definitions and are not directly related to a thermodynamic entropy. The replica correlation measures could be improved by incorporating information about the assigned "overlapping" configurations (nodes may belong to more than one cluster) such as in ref. [95], by modifying the measures to account for the fact that same-type atoms are indistinguishable in this type of a model, or by utilizing a thermodynamic entropy since we are dealing with a physical system in the current application.

Appendix B. MRO and structure factor prepeaks

One experimental approach to ascertain MRO is to look for "prepeaks" in the scattering data. That is, one can look for lower amplitude peaks in the structure factor $S(q)$ which precede the dominant $S(q)$ peak for wave numbers q. While the approach may capture general MRO structures, it is possible to have MRO structures *without significant prepeak(s)* in the structure factor plot.

We illustrate the basic premise of this statement with an elementary example —that of a random arrangement of crystallites. Even though the Fourier transform of the mass density in each individual crystallite has sharp peaks at the reciprocal lattice vectors corresponding to these small crystallites, the structure factor obtained from the entire system may have those peaks vanish. We may, for instance, denote the locations of the centers of mass of individual grains i by $\boldsymbol{R}_i$ and denote the location of individual atoms j in each grain with respect to its center of mass by $\boldsymbol{r}_j$. In that case, the structure factor is

$$S(\boldsymbol{k}) = \sum_{\boldsymbol{R}_i}\sum_{\boldsymbol{r_j}\in i} \exp\left[i\boldsymbol{k}\cdot(\boldsymbol{R}_i+\boldsymbol{r}_j)\right]. \tag{B.1}$$

Within each individual grain on its own,

$$S_i(\boldsymbol{k}) = \sum_{\boldsymbol{r}_j\in i} \exp\left[i\boldsymbol{k}\cdot(\boldsymbol{R}_i+\boldsymbol{r}_j)\right] \tag{B.2}$$

will be sharply peaked about the corresponding reciprocal lattice vectors of grain i. However, the complete sum,

$$S(\boldsymbol{k}) = \sum_{\boldsymbol{R}_i} \exp\left(i\boldsymbol{k}\cdot\boldsymbol{R}_i\right) S_i(\boldsymbol{k}), \tag{B.3}$$

may vanish if the relative distances between the different crystallites are random and lead to phases $\exp(i\boldsymbol{k}\cdot\boldsymbol{R}_i)$ which obliterate the signatures of order in the individual $S_i(\boldsymbol{k})$.

Appendix C. Overlapping nodes between different communities

We wish to account for the possibility of a given atom being connected to more than one physical cluster. For example, in a cubic lattice, each atom participates in the local structure of multiple unit cells. In community detection, this corresponds to allowing "overlapping" community memberships where a node can be a member of more than one community. To accomplish this task, we select the lowest energy partition at the best resolution(s) of the model network (*i.e.*, value(s) of γ in eq. (3) corresponding to extrema in I_N or V).

First, we fix the initial node memberships including the number of communities q. We then sequentially iterate through the community memberships of each node and make changes according to the following: 1) place the node in any additional (non-member) clusters to which it is bound (a negative energy contribution), or 2) remove the node from any member clusters (except for the original membership) in which the current net energy contribution is positive. This process iterates through all nodes as many times as necessary until no node additions or removals are found. The total computational cost is slightly higher than the initial partitioning cost in sect. 4.2 due to the multiple allowed memberships. See also [95] for another method that allows overlapping multiscale network analysis in a general fashion.

Appendix D. Potential shift in the ternary metallic glass

In fig. 16, we show the NMI correlations for a range of potential shifts ϕ_0 and the Potts model weights γ using $s = 12$ replicas and $t = 10$ trials for all data. The NMI peaks are roughly constant over a range of ϕ_0 for the ternary model glass up to $\phi_0 = 0.1$ (the A-A interaction minimum $\phi_{\min} \simeq 0.24\,\mathrm{eV}$).

Appendix E. Time correlations in the ternary model glass

In fig. 17, we show the NMI correlations for a range of separation times t_s between replicas (units are in MD time steps) and the Potts model weights γ using $s = 12$ replicas and $t = 10$ trials for all data. This plot intuitively shows that the correlations weaken as we increase the separation time between replicas. Through varying the time step between replicas, we examine the correlations as a function of t_s and γ which allows us to determine the most important time scale(s) (or relative equivalence). See fig. 7 for

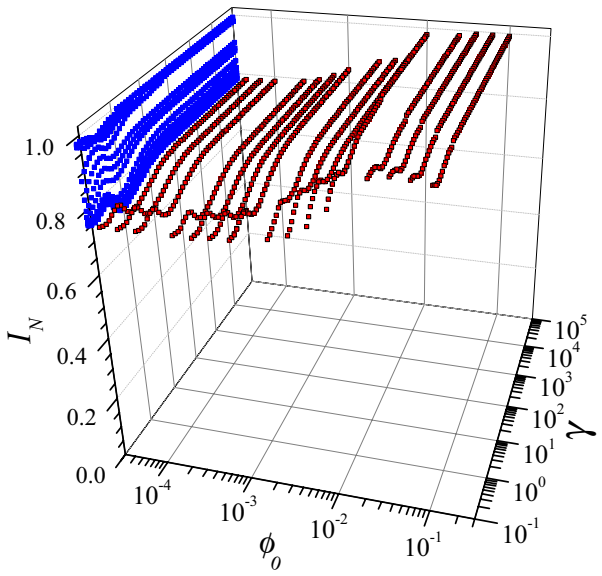

Fig. 16. Plot of NMI I_N as a function of the potential shift ϕ_0 and the Potts model weight γ for the ternary model system in sect. 5.1. The temperature is $T = 300\,\mathrm{K}$ which is *below* the glass transition temperature for this system. This plot shows that the peak is roughly constant across a range of potential shifts. In general systems, we can obtain all natural scales by shifting both ϕ_0 and γ when looking for extrema (or plateaus in some cases). See fig. 17 for the corresponding 2D plot using $\phi_0 = \phi_{\mathrm{avg}}$.

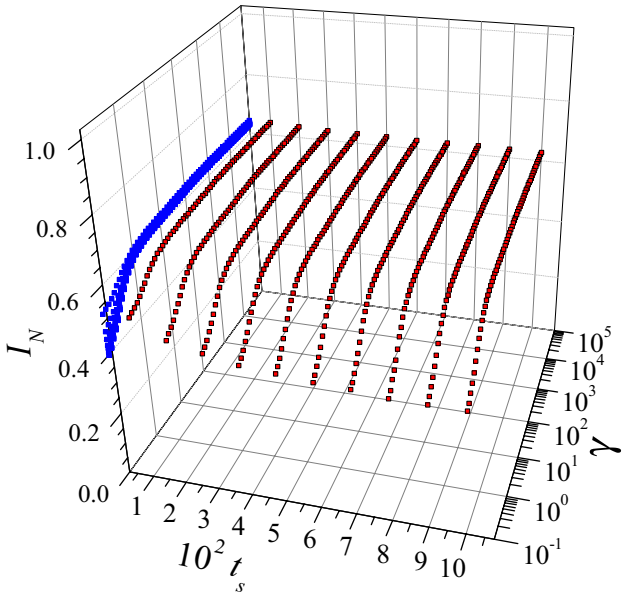

Fig. 17. Plot of NMI I_N as a function of the time (MD time steps) between configurations t_s and the Potts model weight γ for the ternary model system in sect. 5.1. The temperature is $T = 1500\,\mathrm{K}$ which is *above* the glass transition temperature for this system. Intuitively, this plot shows that the correlations become weaker as the time between configurations is increased. This process of examining the correlations as a function of t_s and γ allows us to examine the relevant time scales in addition to the natural spatial scales that we identify in sect. 5.

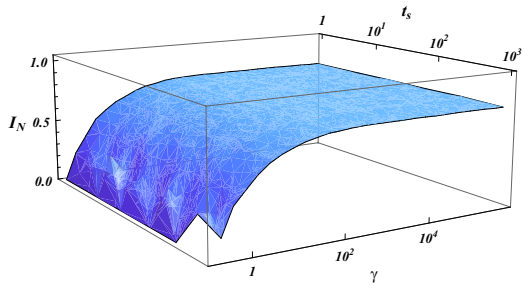

Fig. 18. Plot of NMI I_N as a function of the time (units of 10 MD time steps) between configurations t_s and the Potts model weight γ for the ternary model system in sect. 5.1. The temperature is $T = 0.01$ which is *above* the glass transition temperature for this system. Intuitively, this plot shows that the correlations become weaker as the time between configurations is increased.

the corresponding 2D plot using $t_s = 1000$, and the static limiting case is shown explicitly in appendix J.

Appendix F. Time correlations in the binary LJ glass

In fig. 18, we show the NMI correlations for a range of separation times t_s between replicas (units are in MD time steps) and the Potts model weights γ using $s = 12$ replicas and $t = 10$ trials for all data. This plot shows that the correlations weaken only a little as we increase the separation time between replicas up to 10000 MD time steps. Through varying the time step between replicas, we examine the correlations as a function of t_s and γ which allows us to determine the most important time scale(s) (or relative equivalence). See fig. 7 for the corresponding 2D plot using $t = 1000$, and the static limiting case is shown explicitly in appendix J.

Appendix G. An optimization effect in the binary LJ glass

In fig. 19, we show the NMI correlations for a range of optimization trials p for $s = 12$ replicas and the Potts model weights γ. The plots show that the number of trials p has a small effect on the overall accuracy of the solution for either temperature ($T = 0.01$ or 5), but the effect is slightly more pronounced in the higher temperature $T = 5$ case. See figs. 12 and 14 for the corresponding 2D plots using $s = 10$.

The high temperature clusters are generally much more dispersed. See fig. 15 for sample clusters using $t = 10$ and 20 at $T = 5$ where the clusters correspond to the multiresolution plot in fig. 13 at $\gamma = 10^4$. The corresponding low temperature clusters are analyzed fig. 12 and presented in fig. 14. Panels (a) and (b) in fig. 15 display the typical case of dispersed clusters at $t = 10$. In some instances, the high temperature clusters can be more compact, albeit not densely packed, where panels (c) and (d)

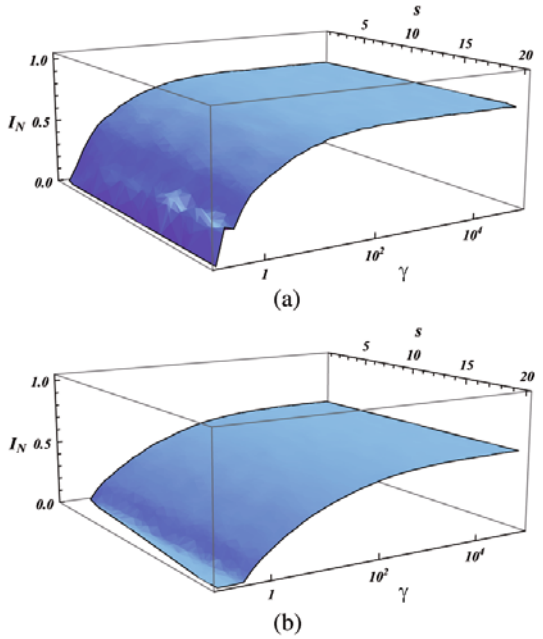

Fig. 19. Plot of NMI I_N as a function of the number of trials s used to optimize the solution and the Potts model weight γ for the binary LJ system in sect. 5.1. The temperatures are $T = 0.01$ in panel (a) and $T = 5$ in panel (b). The number of trials has a higher effect on the accuracy in the $T = 5$ system, but the number of trials does not result in a drastic improvement to the accuracy of our algorithm for either LJ system.

show two examples at $s = 20$. The clusters that are identified are generally consistent in terms of sparseness across all clusters in the solution. The frequency of occurrence of the more compact high temperature clusters, and its dependence on the level of optimization, is a subject for further study.

Appendix H. Multiresolution VI maxima in the LJ system

In fig. 12, we show the multiresolution correlations as a function of the Potts model weight γ for a LJ system of $N = 2000$ nodes using $s = 12$ replicas and $t = 10$ trials for all data. The value of γ corresponding to the peak in V corresponds to the "maximum complexity" of the energy landscape which is often correlated to the system size (see appendix I). Figure 20 shows a sample of the best clusters corresponding to the VI peak which are roughly 300 atoms in size (7–8 atoms in diameter) which is approximately 1/7 of the size of the system. Figure 21 shows some corresponding high temperature $T = 5$ clusters which we note are approximately twice the size ($n \simeq 600$ nodes) and more dispersed.

As an additional note, we remark that, intuitively, one would expect that the VI maxima are directly related to the NMI minimum in this region of γ. This is not the case in this region because the VI metric tends to zero

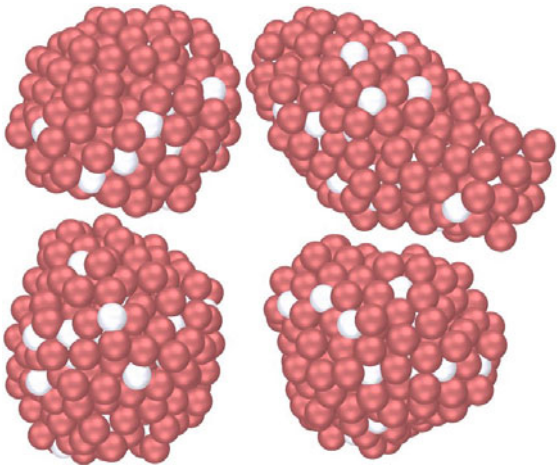

Fig. 20. (Color online) Some of the best clusters for the peak V for the low temperature system shown in fig. 12 in panel (b). The peak in the variation of information V generally correlates with the scale on which the fluctuations in the system division are most prominent. These clusters include overlapping node membership assignments where each node is required to have an overall negative binding energy to the other nodes in the cluster. The atomic identities are B (silver) and A (red) in order of increasing diameters. The resulting configurations constitute tightly bound objects. Representative corresponding high temperature clusters are shown in fig. 21.

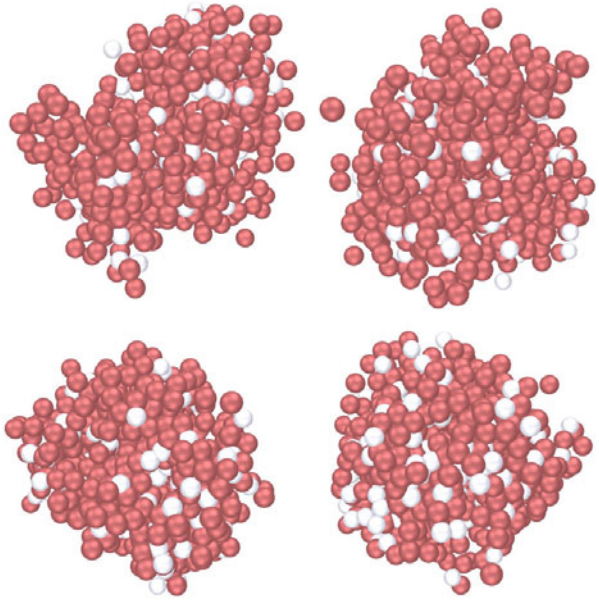

Fig. 21. (Color online) The best clusters found for the peak V for the replica correlation within the high temperature LJ system (see text). These clusters include overlapping node membership assignments where each node is required to have an overall negative binding energy to the other nodes in the cluster. The peak in the variation of information V generally correlates with the scale on which fluctuations are the largest. The atomic identities are B (silver) and A (red) in order of increasing diameters. The diffuse objects found at high temperatures are no longer as compact as at lower temperatures. The normalized mutual information at the peak V is also correspondingly lower than that for the low temperature system.

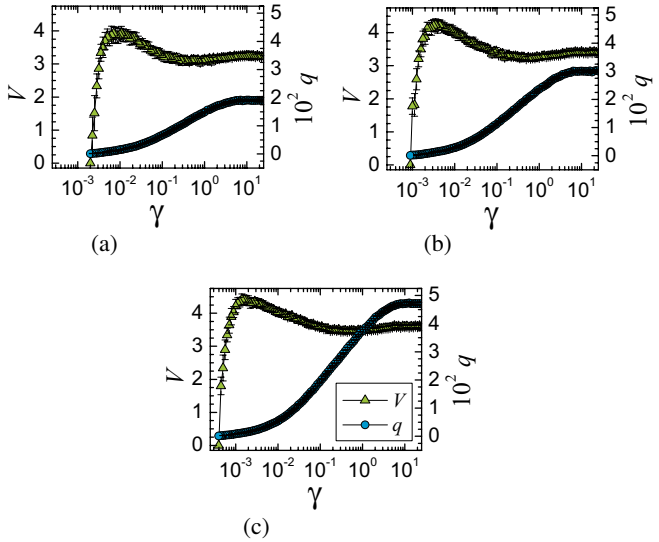

Fig. 22. Plots of the variation of information V between replica pairs as a function of the resolution parameter γ for several system sizes in a LJ simulation (see sect. 3.2). Panels (a), (b), and (c) use $N = 2000$, $N = 4000$, and $N = 8000$, respectively. Note that the value of γ corresponding to the peak V scales downward with the system size (larger structures) which may indicate that it is correlated to a diverging length scale.

as the network collapses to one cluster. It is rapidly shifting from a region of maximum complexity at a peak VI value, where most of the information comes from the sum of the Shannon entropies $H(A)$ and $H(B)$, for two partitions A and B, to minimum complexity at a value of $V = 0$. NMI has a different behavior here, as the ratio $I_N = 2I(A,B)/[H(A) + H(B)]$, because the mutual information between different replicas becomes smaller and approaches zero as the system collapses into larger communities. The low temperature LJ system in fig. 12 shows a transitional very low peak in NMI which corresponds to a near bisection of the system. However, in this case the overlapping cluster configurations collapse to the entire system.

Appendix I. Finite size effects in the LJ system

In fig. 22, we show the VI correlations as a function of the Potts model weight γ for several system sizes of $N = 2000$, $N = 4000$, and $N = 8000$ in panels (a), (b), and (c), respectively. We also used $s = 12$ replicas and $t = 10$ trials for all data. The value of γ corresponding to the peak in V scales downward with the system size (larger structures) indicating that the maximum fluctuations are scaling with the size of the simulation box. In the limit of an infinite size simulation box, the VI peak would tend to $\gamma = 0$.

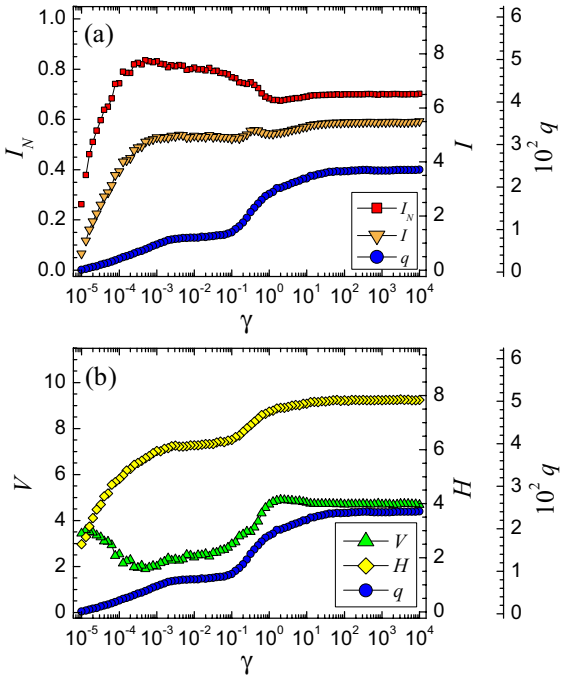

Fig. 23. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). The ternary model system contains 1600 atoms in a mixture of 88% type A, 7% of type B, and 5% of type C with a simulation temperature of $T = 300\,\mathrm{K}$ which is well *below* the glass transition for this system. In this system, we use a single *static* "snapshot" of the system to analyze what our multiresolution algorithm finds at $T = 300\,\mathrm{K}$. This low temperature case shows a preferred resolution at low γ as evidenced by the information extrema at (i) in both panels. See fig. 24 for the corresponding high temperature case.

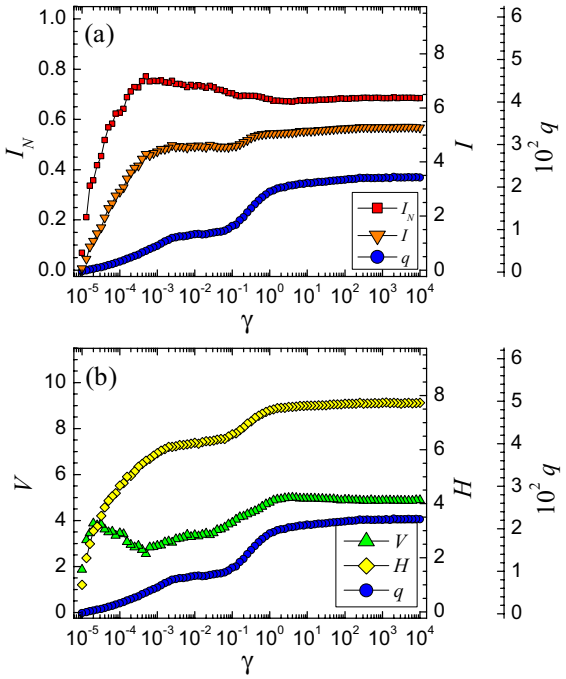

Fig. 24. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). The ternary model system contains 1600 atoms in a mixture of 88% type A, 7% of type B, and 5% of type C with a simulation temperature of $T = 1500\,\mathrm{K}$ which is well *above* the glass transition for this system. In this system, we use a single *static* "snapshot" of the system to analyze what our multiresolution algorithm finds at $T = 1500\,\mathrm{K}$. This high temperature case does show a preferred resolution at low γ as evidenced by the information extrema at (i) in both panels, but the NMI maxima are somewhat lower. The solutions also take longer to solve accurately, meaning that the community detection energy landscape is more complicated. See fig. 23 for the corresponding low temperature case.

Appendix J. Static multiresolution analysis on ternary model glass

The inter-replica NMI/VI correlations change with the time separation interval between the replicas where the longer time separations intuitively result in poorer correlations. The other limiting case is for a *static* analysis of the system. Thus, we are interested in determining what our multiresolution analysis finds when we examine a single-time snapshot of the system. A single snapshot may, potentially, capture a transient feature of the system. This is illustrated in panel (a) of fig. 4 in which the time separation between different replicas is set to zero. We applied the same algorithm as in sect. 4 except that all replicas correspond to the same system time, and we solved the systems with $s = 8$ replicas and $t = 4$ trials per replica. In figs. 23 and 24, our analysis identifies a *static length* associated with the instantaneous configuration. At both high and low temperatures, the static multiresolution analysis displays information extrema at low γ (large spatial scales). The higher temperature system displays weaker correlations with sparser structures whereas the structures of the low temperature system display more significant correlations (and more compact structures). As seen in figs. 23 and 24 the high temperature system displays weaker correlations in the vicinity of the peak NMI than those of the low temperature system. That is, the high temperature solutions require more effort to solve accurately (there are far more metastable minima and more trials are required in order to find better contending solutions). (A similar occurrence was found for the high temperature LJ system in fig. 15.) The solved clusters (not depicted) are comparable to those identified for the time-separated replicas in sect. 5 for both system temperatures.

Appendix K. Static multiresolution analysis on LJ glass

As in appendix J for the ternary system, we further tested a static version of the binary LJ model where the results are similar to the dynamic case in sect. 5.2 both in the MRA plot and the resulting clusters. A notable exception

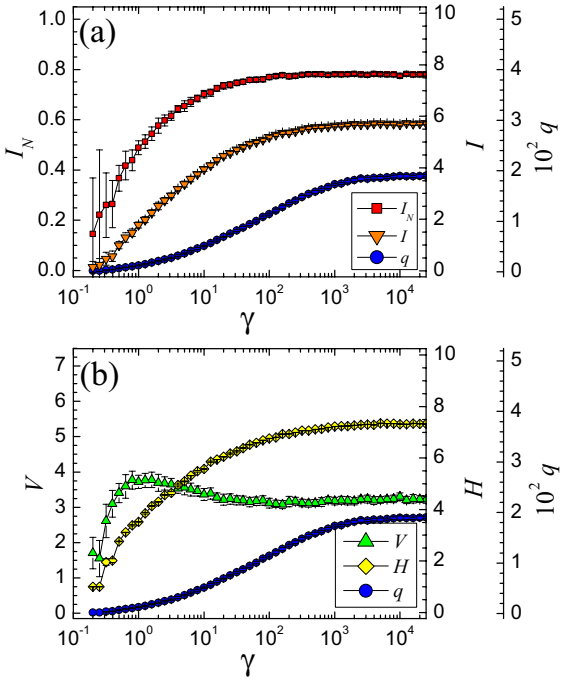

Fig. 25. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). As in sect. 5.2, this LJ system contains 2000 atoms in a mixture of 80% type A and 20% type B with a simulation temperature of $T = 0.01$ (energy units) which is well *below* the glass transition for this system. In this test, we use a single *static* "snapshot" of the system. This low temperature system has a higher overall level of correlation among the replicas than the high temperature case in fig. 26. The overall results are very similar to those observed for the time separated replicas in fig. 12. See fig. 27 for a sample of the best observed clusters.

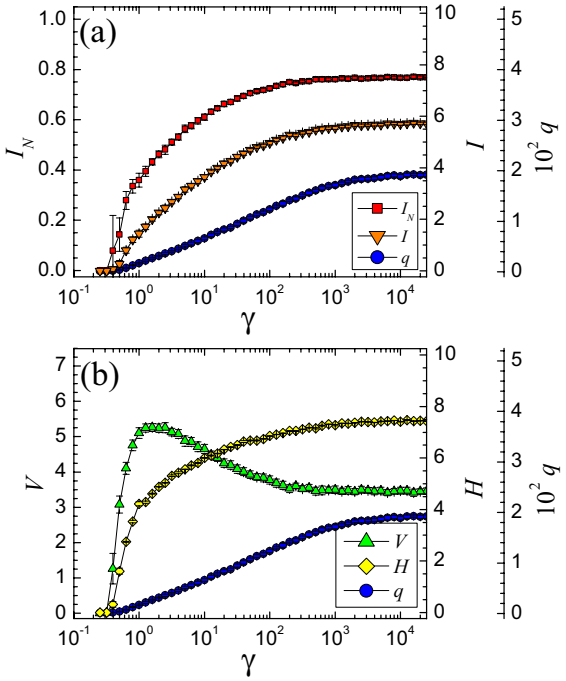

Fig. 26. Panels (a) and (b) show the plots of information measures I_N, V, H, and I and the number of clusters q (right-offset axes) *versus* the Potts model weight γ in eq. (3). The LJ system contains 2000 atoms in a mixture of 80% type A and 20% type B with a simulation temperature of $T = 5$ (energy units) which is well *above* the glass transition for this system. In this analysis, we use a single *static* "snapshot" of the system. This high temperature system has a lower overall correlation than the low temperature case in fig. 25, and the overall results are very similar to what was observed for the time separated replicas in fig. 13 of sect. 5.2.

is that the distinction between the high and low temperature cases becomes more subtle in the multiresolution plots.

We applied the same algorithm as in sect. 4 except that all replicas correspond to the same system time, and we solved the systems with $s = 12$ replicas and $t = 10$ trials per replica. In figs. 25 and 26, we show our multiresolution analysis associated with the instantaneous configurations. These two static plots are similar, but the high temperature case has slightly weaker correlations (lower NMI peak and higher VI minimum). In figs. 25 and 26 where we use time-separated replicas, the contrast in the MRA plots at different temperatures is stronger. Figure 27 shows a sample of the best clusters in the low temperature case which are comparable to those identified for the time-separated replicas in fig. 14 in sect. 5.2.

Appendix L. Multiresolution analysis via measured pair correlation function

As stated in sect. 2, we may apply the same multiresolution network clustering ideas for other structure mod-

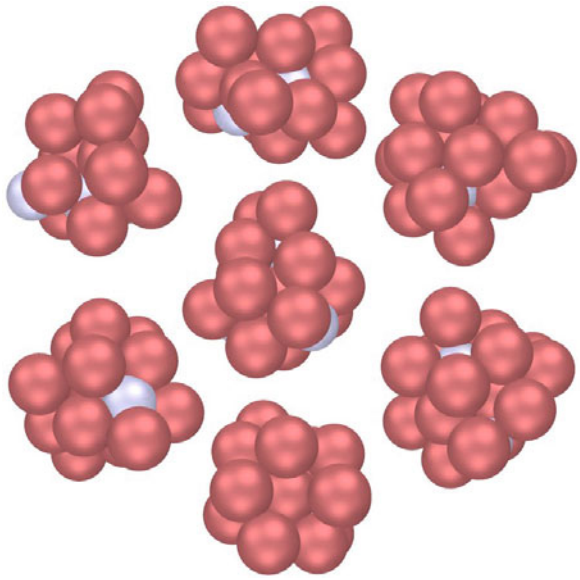

Fig. 27. (Color online) A depiction of some of the best clusters for the peak replica correlation in the static system of fig. 25. These clusters include overlapping node membership assignments where each node is required to have an overall negative binding energy to the other nodes in the cluster. The atomic identities are A (red) and B (silver).

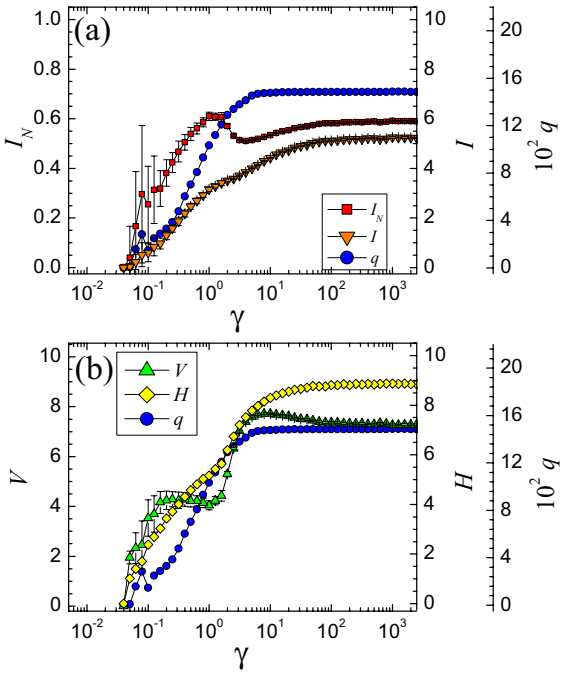

Fig. 28. A plot for the multiresolution analysis of a binary $Zr_{80}Pt_{20}$ system at 250 K below the liquidus. Atomic configurations were generated using conventional RMC methods that are consistent with the experimentally determined scattering data for liquid $Zr_{80}Pt_{20}$ at 250 K below the liquidus. The plot shows a poorly correlated, but nevertheless well-defined, peak in NMI near $\gamma \simeq 1$.

els. In particular, in fig. 28, we apply the method to experimentally adduced pair correlation functions $g_{ab}(r)$ in an amorphous ZrPt system between the different of the different components a and b (*i.e.*, Zr-Zr, Zr-Pt, and Pt-Pt). That is, we set in eq. (3), $A_{ij} = -(g_{ij} + \overline{g}_{ij})$ if $(g_{ij} + \overline{g}_{ij}) < 0$ and $B_{ij} = (g_{ij} - \overline{g}_{ij})$ if $(g_{ij} - \overline{g}_{ij}) > 0$. Here, g_{ij} denotes the pair correlation between the atoms corresponding to nodes i and j, $\overline{g}_{ij}$ is a background average (which we set to zero). Some sample clusters then found by our method are seen in fig. 29. Due to the large system size, we used $s = 8$ replicas and $t = 4$ trials per replica.

Amorphous systems with extended atomic order beyond the nearest neighbor shell provide an excellent framework to test the identification of natural structural elements. The ZrPt is a system that has been shown to have MRO in both the glassy and liquid state [107–111]. MRO, or correlations beyond direct chemical bonding manifest as pre-peaks in the static structure factor extracted from scattering studies of liquids and glasses. The dominant structural elements at the heart of such an ordering are of extreme interest. Atomic configurations that are consistent with the experimentally determined scattering data for liquid $Zr_{80}Pt_{20}$ at 250 K below the liquidus were generated using conventional Reverse Monte Carlo (RMC) methods [112–114]. The result of analysis uses unconstrained RMC which may or may not emulate the precise microscopic structure. We then analyzed a representative

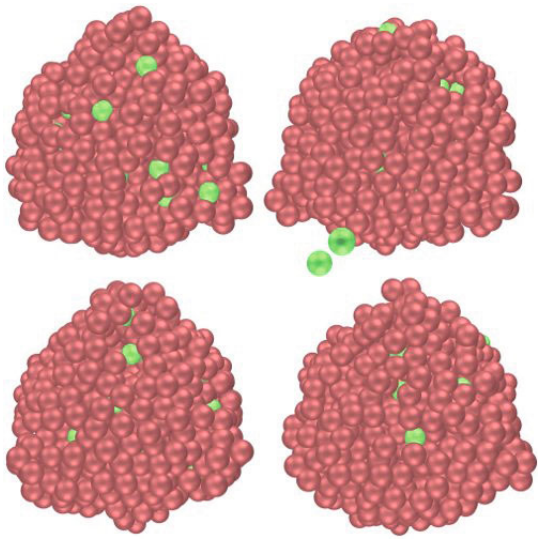

Fig. 29. (Color online) A sample of clusters determined by a multiresolution clustering analysis based on the pair correlation function in fig. 28. The system is a $Zr_{80}Pt_{20}$ system at 250 K below the liquidus (see text) where Zr atoms are depicted as red, and Pt atoms are depicted as green. Each cluster has approximately 700–800 atoms.

system with $N = 7500$ nodes using the algorithm outlined in sect. 4.5.

Specifically, we analyze the system as a *static* model. That is, each replica is based on the one particular system representation since the data is obtained by RMC methods rather than by a dynamical simulation. The multiresolution analysis (MRA) is seen in fig. 28. A few of the relatively uniform large clusters are shown in fig. 29 where each cluster is approximately 700–800 atoms in size. Interestingly, the ZrPt system shows a well-defined secondary MRA NMI peak near $\gamma \simeq 1$. Although the peak NMI value has a poor overall correlation among the replica solutions (low I_N), it is notable that the system displays this secondary peak since it is entirely absent in the LJ binary liquid results in sect. 5.2.

Appendix M. Multiresolution analysis in space-time

In addition to the potential interaction energies, we may model our multiresolution community detection analysis using the mechanical action

$$S = \frac{1}{2}\int \mathrm{d}t \left[\sum_{i=1}^{N} M_i \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t}\right)^2 - \frac{1}{2}\sum_{i\neq j}\phi(\boldsymbol{r}_i, \boldsymbol{r}_j)\right]. \quad \text{(M.1)}$$

Here, M_i is the mass of particle i and $\boldsymbol{r}_i(t)$ denotes its location as a function of time. As throughout, ϕ is the two-particle interaction. Below, we refer to the case of a system in D spatial dimensions. In a discretized form (allowing, in particular, for discrete time steps Δt along

the time axis), the action is

$$S = \sum_{ab} W_{ab}, \tag{M.2}$$

where a and b represent nodes in space-time and W_{ab} is an effective "interaction weight" between them. Thus, we represent our physical system as a network embedded in $(D+1)$ dimensional space-time. We extremize this action "S" corresponding to the Hamiltonian of eq. (3). The spatial links are determined by ϕ_{ij} similar to before (apart from a sign inversion) while the links parallel to the time axis are determined by the particle masses. As before, extrema of the information theory correlation pinpoint the natural structures and scales in the system. In this case, there is only one tuning parameter γ as applied to different replicas of complete space-time networks. Thus, the space and time scales cannot be independently adjusted and are intertwined, being related by a ratio which is the typical speed of sound c (or an appropriate average of such speeds) in equilibrated solids.

That is, if we expand to harmonic order the energy about any of its nearly local minima and denote $\boldsymbol{u}_i$ as the displacement of atom i about an equilibrium position, then we can express the action of the system as an energy of an anisotropic elastic solid in space-time [31,115],

$$S = \frac{1}{2} \int \mathrm{d}^D x \int \mathrm{d}t \, \partial_\mu u^\alpha C_{\mu\nu\alpha\beta}(x,t) \partial_\beta u^\nu, \tag{M.3}$$

with $C_{\mu\nu\alpha\beta}$ being an effective elastic moduli for which amorphous media are generally space (and time) dependent. Here, μ and ν are space-time indices ($\mu = 0, 1, \ldots, D$ for an elastic solid in D spatial dimensions and the zeroth time direction) and $\alpha, \beta = 1, 2, \ldots, D$ are the spatial Cartesian components. Along the time axis, $C_{\mu=0,\nu=0,\alpha\alpha}$ is the mass density $\rho(x,t)$: the effective elastic constant in space-time $C_{\mu=0,\nu=0,\alpha\alpha} = \rho(x,t)$. When μ and ν both assume values between 1 and D, the quantities $C_{\mu\nu\alpha\beta}$ are the elastic moduli of the solid (*e.g.*, the shear modulus μ). The square root of the typical ratio between the usual spatial elastic moduli and those along the time axis for the space-time representation of the solid in eq. (M.3) is given by a typical speed of sound in the system $c = \mathcal{O}(\sqrt{\mu/\rho})$.

Appendix N. Continuum elasticity about local minima and the shear penetration depth

The focus of our article is on a detailed *bona fide* description of amorphous materials that invokes graph theory methods to ascertain general structure. The various examples shown have hopefully clearly outlined the strengths of our approach. In this appendix, we go beyond our earlier atomistic treatment and outline a continuum formalism. Similar to our replica inspired approach thus far, we may consider multiple copies of the entire system wherein, at low temperatures, the atomic configuration (in each copy of the system) veers towards a local energy minimum. In what follows, we consider (continuum limit) fluctuations about these low energy amorphous states and discuss how, generally, rigidity can be associated with a diverging length scale.

Continuum elasticity in regular solids does not assume detailed knowledge of the underlying atomic constituents. In this section, we follow suit and outline general considerations for amorphous solids. In an ideal crystalline system, any small deformation about the crystalline ground state will raise the energy in a harmonic manner. The increase of energy (for non-plastic deformations) is captured by the elastic constants of the material. Nothing prevents, in principle, the application of these considerations for deformation about local energy minima —as these pertain to an amorphous solid in a local energy minimum (an *inherent* structure as it would have pertained to supercooled liquids). In what briefly follows, we consider the harmonic expansion of eq. (M.2) about such a local energy minimum state. We then review general considerations about "shear penetration depth" —the length scale on which the medium reacts to shear. Earlier treatments considered the shear penetration depth in homogeneous media [115, 116]. Within a field theoretic formalism, this length scale stems from a Higgs effect for gauge fields that are associated with elastic stress. Global translational invariance of regular solids implies that, in the long wavelength limit, acoustic sound waves (or phonons) have frequencies that tend to zero (and constitute "Goldstone modes"). As in other arenas (*e.g.*, the Standard Model of particle physics or superconductors), a Higgs effect associated with the minimal coupling of gauge fields to such degrees of freedom, can, when symmetry is broken and a condensate is formed, lead to a screening of mediated interactions. Here, we consider what may occur when replicating these considerations for deformations about a general non-uniform local ground state.

In what follows, we label the displacement of the i-th atom, $\boldsymbol{u}_i$ about its position in a given state, $\boldsymbol{R}_i$ by

$$\boldsymbol{u}_i = \boldsymbol{r}_i - \boldsymbol{R}_i. \tag{N.1}$$

In eq. (N.1), $\boldsymbol{r}_i$ denotes the location of the displaced atoms. Henceforth, we consider a coarse-grained description in which $\boldsymbol{u}$ is a field defined in the continuum. (That is, we replace the lattice indices i in eq. (N.1) by continuous spatial coordinates x.) The displacement vector $\boldsymbol{u}$ is a D-dimensional vector in real space. We apply field theoretic ideas introduced by [117]. In particular, we next follow the procedure of [31,115,116] and write the corresponding action of the system in a manner similar to that of the energy of an anisotropic solid in $(D+1)$ dimensions. We will label the time direction as the $\alpha = 0$ direction and denote all spatial Cartesian coordinates by $\alpha = 1, 2, \ldots, D$. We use Latin indices $a = 1, 2, \ldots, D$ to label the spatial directions alone. In the continuum limit, expanding the energy to harmonic order about any local minima (inherent state), the action reads [115,31],

$$S = \frac{1}{2} \int \mathrm{d}^D x \int \mathrm{d}t \, \partial^\alpha u^a C_{\alpha\beta ab}(x,t) \partial^\beta u^b. \tag{N.2}$$

In eq. (N.2), repeated indices are to be summed over. The quantities $C_{\alpha\beta ab}$ derive from effective elastic moduli. In regular solids, $C_{\alpha\beta ab}$ are space (and time) independent quantities. In an amorphous medium, however, $C_{\alpha\beta ab}(x,t)$ are generally *space (and time) dependent*. The effective elastic constant C_{00ab} in eq. (N.2) is set by the mass density, *i.e.*, $C_{00ab} = \rho(x,t)$ [31,115,116]. When α and β both assume values between 1 and D, the quantities $C_{\alpha\beta ab}$ are set by the elastic moduli of the solid (*e.g.*, the shear modulus μ). The square root of the typical ratio between the usual spatial elastic moduli and those along the time axis for the space-time representation of the solid in eq. (N.2) is given by a typical speed of sound in the system $c = \mathcal{O}(\sqrt{\mu/\rho})$.

One approach to defining defects such as dislocation and disclinations in regular solids is to consider the net change in the displacement or local orientation associated with closed trajectories. Defect densities relative to a given inherent state may be defined in a similar way. More formally, we can dualize (via a Hubbard Stratonovich transformation) the quadratic action of eq. (N.2) and associate the dual fields with defect current densities in space-time. Replicating the analysis performed for regular solids in [31,115,116] to amorphous solids with an invertible space-time dependent elastic moduli $C_{\alpha\beta ab}$, we obtain the same result obtained for regular solids. Namely, dislocation and disclination current densities are given by

$$\begin{aligned} J^a_{\alpha\beta} &= \epsilon_{\alpha\beta\lambda\rho}\partial^\lambda\partial^\rho u^a, \\ T^{\alpha\beta}_{\gamma\delta} &= \epsilon_{\gamma\delta\lambda\rho}\partial^\lambda\partial^\rho \omega^{\alpha\beta}, \end{aligned} \tag{N.3}$$

with the local rotation given by

$$\omega_{\alpha\beta} = \frac{1}{2}\epsilon_{\alpha\beta\lambda a}\partial_\lambda u^a. \tag{N.4}$$

The temporal direction ($\alpha = 0$) components of these tensors are the usual (dislocation and disclination) defect densities of elasticity,

$$\begin{aligned} \alpha^a_i &= \epsilon_{ijk}\partial^j\partial^k u^a, \\ \Theta^a_i &= \frac{1}{2}\epsilon_{ijk}\epsilon^{abc}\partial^j\partial^k\partial_b u^c. \end{aligned} \tag{N.5}$$

The condition of a divergence-less stress tensor (σ) in the absence of an applied force enables us to write it as a curl of a gauge field

$$\sigma^a_\alpha = \epsilon_{\alpha\beta\lambda}\partial^\beta B^{\lambda a}. \tag{N.6}$$

Elastic shear is mediated by the gauge field B that mediates "shear photons" [115,116]. When a plastic component of the displacement field $\boldsymbol{u}$ exists, it couples minimally to the stress-gauge fields via

$$S_{\rm disl} = \int \mathrm{d}^D x\, \mathrm{d}t\, B^{\alpha a} J_{\alpha a} \tag{N.7}$$

for the coupling between the dislocation currents and gauges. Mathematically, the structure is very similar to that in electromagnetism when currents couple minimally to gauge fields. In the Higgs borne effect that results, just as charges screen applied electromagnetic fields (as, *e.g.*, in the Debye screening of plasmas), defect charges can screen the elastic shear. In the uniform medium, when a defect condensate ($|\Psi_0| \neq 0$) appears, it gives rise to a screening of the elastic shear [31,115], with a screening length set by

$$\lambda_{\rm shear} = \frac{c_T}{|\psi_0|\sqrt{\mu}}\,. \tag{N.8}$$

In eq. (N.8), c_T is the transverse sound velocity and μ is the shear modulus. In the case of simple dilation of the elastic moduli when $C_{\alpha\beta ab}(x,t) = w(x,t)K_{\alpha\beta ab}$, where $K_{\alpha\beta}$ is a space-time independent tensor and w is a dilation function, the calculation of the elastic shear penetration depth can be replicated for the amorphous system to yield eq. (N.8) when the quantities refer to the local values of $\lambda_{\rm shear}$, c_T, ψ_0, and μ. A similar effect appears for general space-time dependent elastic moduli [118] when considering elastic deformations about a local energy minimum. As is seen from eq. (N.8), *when the defect density tends to zero, the elastic shear penetration depth diverges*

$$\lambda_{\rm shear} \to \infty, \tag{N.9}$$

as the system becomes rigid throughout. In order to ascertain detailed shear response in a system, we may employ the shear stress correlation functions as link weights within our community detection algorithm. This is topic of a future study.

References

1. P. Ronhovde, S. Chakrabarty, M. Sahu, K.K. Sahu, K.F. Kelton, N. Mauro, Z. Nussinov, `http://arxiv.org/pdf/1101.0008` (2010).
2. R. Zallen, *The Physics of Amorphous Solids* (John Wiley & Sons, Inc., 1983) pp. 23–32.
3. A.L. Greer, E. Ma, MRS Bull. **32**, 611 (2007).
4. B.C. Hancock, M. Parks, Pharma. Res. **17**, 397 (2000).
5. M. Telford, Mater. Today **7**, 36 (2004).
6. M. Wuttig, N. Yamada, Nat. Mater. **6**, 824 (2007).
7. J. Zaanen, Nature **404**, 714 (2000).
8. S.A. Kivelson, I.P. Bindloss, E. Fradkin, V. Oganesyan. J.M. Tranquada, A. Kapitulnik, C. Howald, Rev. Mod. Phys. **75**, 1201 (2003).
9. F.C. Frank, Proc. R. Soc. London, Ser. A. **215**, 43 (1952).
10. T. Schenk, D. Holland-Moritz, V. Simonet, R. Bellissent, D.M. Herlach, Phys. Rev. Lett. **89**, 075507 (2002).
11. K.F. Kelton, G.W. Lee, A.K. Gangopadhyay, R.W. Hyers, T.J. Rathz, J.R. Rogers, M.B. Robinson, D.S. Robinson, Phys. Rev. Lett. **90**, 195504 (2003).
12. D.R. Nelson, *Defects and Geometry in Condensed Matter Physics* (Cambridge University Press, Cambridge, 2002).
13. J.F. Sadoc, R. Mosseri, *Geometrical Frustration* (Cambridge University Press, Cambridge, 1999).
14. G. Tarjus, S.A. Kivelson, Z. Nussinov, P. Viot, J. Phys.: Condens. Matter **17**, R1143 (2005).
15. Z. Nussinov, Phys. Rev. B **69**, 014208 (2004).
16. V. Lubchenko, P.G. Wolynes, Annu. Rev. Phys. Chem. **58**, 235 (2007).

17. L. Angelani, G. Parisi, G. Ruocco, G. Viliani, Phys. Rev. E **61**, 1681 (2000).
18. G. Parisi, Physica A **280**, 115 (2000).
19. S. Sastry, P.G. Debenedetti, F.H. Stillinger, Nature **393**, 554 (1998).
20. P.G. Debenedetti, F.H. Stillinger, Nature **410**, 259 (2001).
21. V. Lunchenko, P.G. Wolynes, J. Chem. Phys. **121**, 2852 (2004).
22. W.L. Johnson, M.D. Demetriou, J.S. Harmon, M.L. Lind, K. Samwer, MRS Bull. **32**, 644 (2007).
23. Jonathan P.K. Doye, David J. Wales, Fredrik H.M. Zetterling, Mikhail Dzugutov, J. Chem. Phys. **118**, 2792 (2008).
24. T.R. Kirkpatrick, D. Thirumalai, P.G. Wolynes, Phys. Rev. A **40**, 1045 (1989).
25. M. Tarzia, M.A. Moore, Phys. Rev. E **75**, 031502 (2007).
26. W. Gotze, J. Phys.: Condens. Matter **11**, A1 (1999).
27. P. Mayer, K. Miyazaki, D.R. Reichman, Phys. Rev. Lett. **97**, 095702 (2006).
28. J.P. Garrahan, J. Phys.: Condens. Matter **14**, 1571 (2002).
29. D. Kivelson, S.A. Kivelson, X. Zhao, Z. Nussinov, G. Tarjus, Physica A **219**, 27 (1995).
30. F. Ritort, P. Sollich, Adv. Phys. **52**, 219 (2003).
31. V. Cvetkovic, Z. Nussinov, J. Zaanen, Philos. Mag. **86**, 2995 (2006).
32. E. Aharonov, E. Bouchbinder, H.G.E. Hentschel, V. Ilyin, N. Makedonska, I. Procaccia, N. Schupper, EPL **77**, 56002 (2007).
33. L. Berthier, G. Biroli, arXiv: 1011.2578 (2010).
34. D. Chandler, J.P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).
35. A. Montanari, G. Semerjian, J. Statis. Phys. **125**, 23 (2006).
36. H. Tanaka, T. Kawasaki, H. Shintani, K. Watanabe, Nat. Mater. **9**, 324 (2010).
37. M. Mosayebi, E.D. Gado, P. Iig, H.C. Ottinger, Phys. Rev. Lett. **104**, 205704 (2010).
38. L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, D. El Masri, D. L'Hote, F. Ladieu, M. Pierno, Science **310**, 1797 (2005).
39. S. Karmakar, C. Dasgupta, S. Sastry, Proc. Natl. Acad. Sci. U.S.A. **106**, 3675 (2010).
40. G. Biroli, J.-P. Bouchaud, A. Cavagna, T.S. Grigera, P. Verrocchio, Nat. Phys. **4**, 771 (2008).
41. G. Biroli, J.-P. Bouchaud, e-print arXiv:0912.2542 (2009).
42. J. Kurchan, D. Levine, e-print arXiv:0904.4850 (2009).
43. J.D. Bernal, Nature **185**, 68 (1960).
44. D.B. Miracle, W.S. Sanders, O.N. Senkov, Philos. Mag. **83**, 2409 (2003).
45. D.B. Miracle, Nat. Mater. **3**, 697 (2004).
46. D.B. Miracle, T. Egami, M.K. Flores, K.F. Kelton, MRS Bull. **32**, 629 (2007).
47. D.B. Miracle, E.A. Lord, S. Ranganathan, Mater. Trans. **47**, 1737 (2006).
48. W.K. Luo, H.W. Sheng, F.M. Alamgir, J.M. Bai, J.H. He, E. Ma, Phys. Rev. Lett. **92**, 145502 (2004).
49. P. Ganesh, M. Widom, Phys. Rev. B **77**, 0145205 (2008).
50. Y.T. Shen, T.H. Kim, A.K. Gangopadhyay, K.F. Kelton, Phys. Rev. Lett. **102**, 057801 (2009).
51. E.D. Gado, P. Ilg, M. Kroger, H.C. Ottinger, Phys. Rev. Lett. **101**, 095501 (2008).
52. W. Kob, C. Donati, S.J. Plimpton, P.H. Poole, S.C. Glotzer, Phys. Rev. Lett. **79**, 2827 (1997).
53. E.R. Weeks, J.C. Crocker, A.C. Levitt, A. Schofield, D.A. Weitz, Science **287**, 627 (2000).
54. A. Widom-Cooper, P. Harrowell, H. Fynewever, Phys. Rev. Lett. **93**, 135701 (2004).
55. J.D. Stevenson, J. Schmalian, P.G. Wolynes, Nat. Phys. **2**, 268 (2006).
56. L. Berthier, G. Biroli, J.-P. Bouchaud, R.L. Jack, e-print: 1009.4665 (2010).
57. H.W. Sheng, W.K. Luo, F.M. Alamgir, J.M. Bai, E. Ma, Nature **439**, 419 (2006).
58. J.L. Finney, Proc. R. Soc. London, Ser. A **319**, 479 (1970).
59. J. Dana Honeycutt, Hans C. Andersen, J. Phys. Chem. **91**, 4950 (1987).
60. P.J. Steinhardt, D.R. Nelson, M. Ronchetti, Phys. Rev. B **28**, 784 (1983).
61. T.C. Hufnagel, S. Brennan, Phys. Rev. B **67**, 014203 (2003).
62. M.M.J. Treacy, J.M. Gibson, L. Fan, D.J. Paterson, I. McNulty, Rep. Prog. Phys. **68**, 2899 (2005).
63. P. Ronhovde, Z. Nussinov, Phys. Rev. E **80**, 016109 (2009).
64. Peter Ronhovde, Zohar Nussinov, Phys. Rev. E **81**, 046114 (2010).
65. M. Girvan, M.E.J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 7821 (2002).
66. S. Fortunato, Phys. Rep. **486**, 75 (2010).
67. A. Lancichinetti, S. Fortunato, Phys. Rev. E **80**, 056117 (2009).
68. M.E.J. Newman, M. Girvan, Phys. Rev. E **69**, 026113 (2004).
69. V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, J. Stat. Mech. **10**, P10008 (2008).
70. M.E.J. Newman, Phys. Rev. E **69**, 066133 (2004).
71. J. Reichardt, S. Bornholdt, Phys. Rev. E **74**, 016110 (2006).
72. M.E.J. Newman, Phys. Rev. E **74**, 036104 (2006).
73. V. Gudkov, V. Montelaegre, S. Nussinov, Z. Nussinov, Phys. Rev. E **78**, 016113 (2008).
74. A. Arenas, A. Díaz-Guilera, C.J. Pérez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).
75. M. Rosvall, C.T. Bergstrom, Proc. Natl. Aca. Sci. U.S.A. **105**, 1118 (2008).
76. Book chapter by U. Brandes, D. Dellng, M. Gaertler, R. Gorke, M. Hoefer, Z. Nikoloski, D. Wagner, *On finding graph clusterings with maximum modularity* in *Graph-Thoeretic Concepts in Computer Science*, *Springer Lect. Notes Comput. Sci.*, DOI: 10.1007/978-3-540-74839-7 (2007).
77. M. Meila, J. Multivariate Anal. **98**, 873 (2007).
78. This includes, *e.g.*, i) relaxation processes of structural glass formers relating to slow global (α) process and local rapid rearrangments (β processes), and ii) systems with multiple degrees of freedom (structural, magnetic, and other) that may exhibit transitions at different temperatures.
79. M. Mezard, A. Montanari, J. Stat. Phys. **124**, 1317 (2006).
80. S. Franz, A. Montanari, J. Phys. A. **40**, F251 (2007).
81. K.K. Sahu, N.A. Mauro, L. Longstreth-Spoor, D. Saha, Z. Nussinov, M.K. Miller, K.F. Kelton, Acta Mater. **58**, 4199 (2010).

82. T. Egami, J. Non-Cryst. Solids **317**, 30 (2003).
83. W. Kob, H.C. Andersen, Phys. Rev. E **51**, 4626 (1995).
84. L.C. Valdes, F. Affouard, M. Descamps, J. Habasaki, J. Chem. Phys. **130**, 154505 (2009).
85. J. Stadler, R. Mikula, H.R. Trebin, Int. J. Mod. Phys. C **8**, 1131 (1997).
86. M. Mihalkovič, C.L. Henley, M. Widom, P. Ganesh, e-print arXiv:cond-mat.mtrl-sci/0802.2926 (2008).
87. VASP website: `http://cms.mpi.univie.ac.at/vasp/`.
88. J.A. Moriarty, M. Widom, Phys. Rev. B **56**, 7905 (1997).
89. For overdamped systems (such as particles in viscous liquids) obeying "Aristotelian dynamics" ($\boldsymbol{F} \propto \boldsymbol{v}$), clusters in which the force on all their individual particles is nearly the same will move with uniform velocity. This leads to high velocity correlations between particles belonging to the same cluster.
90. D. Hu, P. Ronhovde, Z. Nussinov, e-print arXiv: 1008.2699, to be pubjished in Philos. Mag. (2011).
91. S. Mossa, G. Tarjus, J. Chem. Phys. **119**, 8069 (2003).
92. P. Ronhovde, S. Chakrabarty, M. Sahu, K.K. Sahu, K.F. Kelton, N. Mauro, Z. Nussinov, e-print arXiv:1102.1519.
93. A. Arenas, A. Fernandez, S. Gomez, New. J. Phys. **10**, 053039 (2008).
94. J.M. Kumpula, J. Saramaki, K. Kaski, J. Kertesz, Fluct. Noise Lett. **7**, L209 (2007).
95. A. Lancichinetti, S. Fortunato, J. Kertesz, New J. Phys. **11**, 033015 (2009).
96. P.J. Mucha, T. Richardson, K. Macon, M.A. Porter, J.-P. Onnella, Science **328**, 876 (2010).
97. D.J. Fenn, M.A. Porter, M. McDonald, S. Wiliams, N.F. Johnson, Chaos **19**, 033119 (2009).
98. D. Hu, P. Ronhovde, Z. Nussinov, e-print arXiv:1106.5793 (2011).
99. Asaph Widmer-Cooper, Heidi Perry, Peter Harrowell, David R. Reichman, Nat. Phys. **4**, 711 (2008).
100. J.X. Lin, C. Reichhardt, Z. Nussinov, L.P. Pryadko, C.J. Olson Reichhardt, Phys. Rev. E **73**, 061401 (2006).
101. M. Lisa Manning, Andrea J. Liu, arXiv:1012.0064 (2010).
102. W. Donath, A. Hoffman, IBM J. Res. Devel. **17**, 420 (1973).
103. M. Fiedler, Czech. Math. J. **23**, 298 (1973).
104. Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence* (Springer-Verlag, Berlin, Germany, 1984).
105. R.B. Griffiths, Phys. Rev. Lett. **23**, 17 (1969).
106. S. Chakrabarty, M. Widom, M. Mihalkovič, K.F. Kelton, Z. Nussinov, in preparation.
107. T. Nakamura, E. Matsubara, M. Sakurai, M. Kasai, A. Inoue, Y. Waseda, J. Non-Cryst. Solids **312-314**, 517 (2002).
108. J. Saida, K. Itoh, K. Sato, M. Imafuku, T. Sanada, A. Inoue, J. Phys.: Condens. Matter **21**, 375104 (2009).
109. D.J. Sordelet, R.T. Ott, M.Z. Li, S.Y. Wang, M.F. Besser, A.C.Y. Liu, M.J. Kramer, Metall. Mater. Trans. A **39**, 1908 (2007).
110. S.Y. Wang, C.Z. Wang, M.Z. Li, L. Huang, R.T. Ott, M.J. Kramer, D.J. Sordelet, K.M. Ho, Phys. Rev. B **78**, 184204 (2008).
111. N.A. Mauro, V. Wessels, J.C. Bendert, S. Klein, A.K. Gangopadhyay, M.J. Kramer, S.G. Hao, G.E. Rustan, A. Kreyssig, A.I. Goldman, K.F. Kelton, Phys. Rev. B **83**, 184109 (2011).
112. R.L. McGreevy, J. Phys.: Condens. Matter **3**, F9 (1991).
113. D.A. Keen, R.L. McGreevy, Nature **344**, 423 (1990).
114. T.H. Kim, K.F. Kelton, J. Chem. Phys. **126**, 054513 (2007).
115. J. Zaanen, Z. Nussinov, S.I. Mukhin, Ann. Phys. (N.Y.) **310**, 181 (2004).
116. V. Cvetkovic, Z. Nussinov, S. Mukhin, J. Zaanen, EPL **81**, 27001 (2008).
117. H. Kleinert, *Gauge Fields in Condensed Matter, Vol. II: Stresses and Defects, Differential Geometry, Crystal Defects* (World Scientific, Singapore, 1989).
118. V. Cvetkovic, M. Vasin, Z. Nussinov, in preparation.