Stability-to-instability transition in the structure of large-scale networks

Dandan Hu, Peter Ronhovde, and Zohar Nussinov

Phys. Rev. E **86**, 066106 — Published 6 December 2012

DOI: 10.1103/PhysRevE.86.066106

The stability to instability transition in the structure of large scale networks

Dandan Hu, Peter Ronhovde, and Zohar Nussinov*
Department of Physics, Washington University in St. Louis,
Campus Box 1105, 1 Brookings Drive, St. Louis, MO 63130, USA

We examine phase transitions between the "easy," "hard," and the "unsolvable" phases when attempting to identify structure in large complex networks ("community detection") in the presence of disorder induced by network "noise" (spurious links that obscure structure), heat bath temperature T, and system size N. The partition of a graph into q optimally disjoint subgraphs or "communities" inherently requires Potts type variables. In earlier work [Phil. Mag. **92**, 406 (2012)] when examining power law and other networks (and general associated Potts models), we illustrated transitions in the computational complexity of the community detection problem typically correspond to spin-glass-type transitions (and transitions to chaotic dynamics in mechanical analogs) at both high and low temperatures and/or noise. When present, transitions at low temperature or low noise correspond to entropy driven (or "order by disorder") annealing effects wherein stability may initially increase as temperature or noise is increased before becoming unsolvable at sufficiently high temperature or noise. Additional transitions between contending viable solutions (such as those at different natural scales) are also possible. Identifying community structure via a dynamical approach where "chaotic-type" transitions were earlier found. The correspondence between the spin-glass-type complexity transitions and transitions into chaos in dynamical analogs might extend to other hard computational problems. In this work, we examine large networks (with a power law distribution in cluster size) that have a large number of communities ($q \gg 1$). We infer that large systems at a constant ratio of q to the number of nodes N asymptotically tend toward insolvability in the limit of large N for any positive T. The asymptotic behavior of temperatures below which structure identification might be possible, $T_\times = O[1/\log q]$, decreases slowly, so for practical system sizes, there remains an accessible, and generally easy, global solvable phase at low temperature. We further employ multivariate Tutte polynomials to show that increasing q emulates increasing T for a general Potts model, leading to a similar stability region at low T. Given the relation between Tutte and Jones polynomials, our results further suggest a link between the above complexity transitions and transitions associated with random knots.

PACS numbers: 89.75.Fb, 64.60.Cn, 89.65.-s

I. INTRODUCTION

Applications of physics to networks [1] has opened fascinating doors for enhancing our understanding of these complex systems. In particular, community detection [2–5] endeavors to identify pertinent structures within such systems. Applications of the problem are exceptionally broad, and numerous methods have been proposed to attack the problem [6–14], some of which have been compared for efficiency and accuracy [15–18].

Computational "phase transitions" have been studied in many challenging problems [19–26]. Practical implications of such studies abound (e.g., Refs. [19, 23, 27–29]), and understanding the behavior of algorithmic solutions to these problems is of interest because the knowledge can be leveraged to understand when a particular solution is computationally challenging, trustworthy, or perhaps not obtainable either via an inherent difficulty or required computational effort. Such knowledge may be used to in certain cases to predict the hard or unsolvable regimes of the problem *a priori* (e.g., k-SAT [20]) or perhaps, more practically in general, to dynamically adapt the solver during the onset of a phase transition [30].

Earlier work related to computational phase transitions with connections to clustering include [31, 32], and Ref. [33] reviewed some critical phenomena in complex networks. The complexity of the energy landscape in community detection was studied for a "fixed" Potts model (model parameters are not set by the network under study) [34, 35], modularity [36], and belief propagation on block models [37]. The former and latter studies explicitly identified phase transitions in the respective systems. We extend a previous analysis [35] of a Potts model where we studied the thermodynamic and complexity character resulting in two distinct transitions: an entropic stabilization transition where added complexity can result in "order by disorder" annealing and a high temperature disordered unsolvable phase. For extreme complexity (high noise) at low T, the system is again unsolvable. Additional transitions can appear between unsolvable and difficult solutions or contending partitions of natural network scales. Here, we seek to move beyond characterizing the solvable/unsolvable transition to study the transitions in terms of changes in the energy landscape and thermodynamic functions as functions of temperature and "noise" (intercommunity edges).

We utilize overlap parameters in the form of information theory measures (see Appendix B) and a "compu-

*Electronic address: zohar@wuphys.wustl.edu

tational susceptibility" χ (see Appendix C). Using these measures, we monitor increases in the number of local minima corresponding to (often sharply) increased computational complexity. We apply our Potts model to solve a random graph with an embedded ground state, and we identify phase transitions between "easy" and "hard" *solvable* phases which transition into *unsolvable* regions. Specifically, the normalized mutual information (NMI) I_N, Shannon entropy H, the energy E, and χ exhibit progressively sharper changes as the system size N increases suggesting the existence of genuine thermodynamic transitions. Similar analysis can be done for other community detection approaches. Many community detection methods will agree on the best solution within the easy phase, but the hard region presents a substantially more difficult challenge.

The identified transitions may be connected to jamming [28, 29] and avalanche (cascade) transitions [38–40] in networks. Dynamic jamming transitions occur in traffic, computer network, particulate matter (e.g., sand piles), and the glassy state in amorphous solids may be caused by similar behavior. Refs. [41–43] showed relations between clustering and cascades in certain networks, and Ref. [44] relates agent dynamics to the Kuramoto oscillators model which has been used for community detection [45]. The threshold emergence of Giant Connected Components (GCC) is related to epidemic thresholds [46, 47], and by nature of the emerging global connectivity, the GCC is directly detectable via clustering at large-scale resolutions [*i.e.*, small γ in Eq. (1)]. Jones polynomials in knot theory are related to Tutte polynomials for the Potts model, so our results suggest similar transitions in random knots (see Appendix G).

We will analytically investigate partition functions and free energies of a several graphs in the high temperature T and large number of communities q approximations. We illustrate that increasing T emulates increasing q for a general system, and the analytical results are consistent with the computational phase diagrams.

The remarks of the paper is organized as follows: We introduce the community detection model in Sec. II and then the embedded graph/noise test in Sec. III. Section IV demonstrates the spin-glass-type transitions that occur in our community detection problem via numerical simulation using several instability measures. In Sec. V, we derive crossover thresholds for a simple case and discuss their connections to the numerical simulations, and Sec. VI demonstrates the effect of the different solution regions with a specific example. Section VII carries out analytic free energy calculations on arbitrary unweighted graphs using a ferromagnetic Potts model. Appendix A exams the notation of "trials" and "replicas" which are of paramount importance in our work to directly probe the phase diagram sans the use of mean-field type or other approximation concerning complexity. Appendix A defines some terminology used in the paper. Appendix B and Appendix C describe our information and stability measures, and Appendix D elaborates on our heat bath community detection algorithm. We introduce the Tutte polynomial method for calculating the partition function of a Potts model for unweighted and weighted graphs in Appendix E, and we show an exact calculation for a simple connected graph in Appendix F. Finally, Appendix G conjectures the existence of a similar transition for knots.

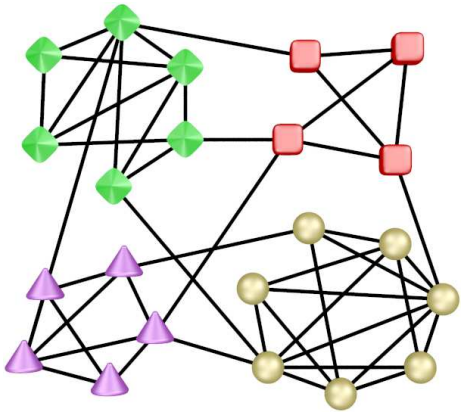


FIG. 1: (Color online) The figure illustrates a partition where nodes are separated into distinct communities as indicated by distinct shapes and colors, thus identifying relevant structure in the graph. The current work elaborates on computational transitions and disorder in terms noise (extraneous intercommunity edges) or thermal effects (high temperature T or large system size N) of solving such systems using a stochastic heat bath solver (see Appendix D).

II. POTTS HAMILTONIAN

We employ a spin-glass-type Potts model Hamiltonian for solving the community detection problem

$$H(\{\sigma\}) = -\frac{1}{2}\sum_{i\neq j}\left[A_{ij} - \gamma\left(1 - A_{ij}\right)\right]\delta(\sigma_i, \sigma_j) \qquad (1)$$

which we refer to as an "Absolute Potts Model" (APM). Given N nodes, A_{ij} denotes the adjacency matrix where $A_{ij} = 1$ if nodes i and j are connected and is 0 otherwise. In general, A_{ij} may be trivially extended to a weighted adjacency matrix w_{ij} (perhaps including "adversarial" relations) [34], but we utilize unweighted graphs in most of the current work (see Sec. VI). Each spin σ_i may assume integer values in the range $1 \leq \sigma_i \leq q$ where q is the (dynamic) number of communities where node i is a member of community k when $\sigma_i = k$. In the current work, we set the resolution parameter [11] to $\gamma = 1$ which is near an optimal value for communities with high internal edge densities (see Sec. III).

Previous work [11, 34] elaborated on a "zero-temperature" ($T = 0$) community detection algorithm which we used to minimize Eq. (1). A depiction of community structure is shown in Fig. 1 where different communities are represented by different node shapes and

colors. Here, we investigate the Hamiltonian at nonzero temperatures ($T > 0$) by applying a heat bath algorithm (HBA, see Appendix D). Briefly, we iteratively select each node and test for possible moves where probabilities are calculated via a Boltzmann weight $e^{-\beta\Delta E}$ ($\beta = 1/T$) at a temperature T using the energy change (ΔE) as if the node were moved into a connected (or new) cluster. Similarly, as elaborated in Appendix D, following each step, we further allow the possibility of community merges based on a Boltzmann weight.

We further invoke s independent solutions ("trials", see Appendix A) by solving copies of the system which differ by a permutation of the order of the spin indices. This process leads to states that (perhaps locally) minimize Eq. (1), so we select the lowest energy trial as the best solution. We vary s in the range $4 \leq s \leq 20$ where we employ $s = 4$ trials in general and use $s > 4$ trials for calculating the computational susceptibility in Eq. (C1). In brief, a "computational susceptibility" monitoring the onset of high complexity, can be defined as $\chi_n = I_N(s = n) - I_N(s = 4)$. χ is the increase in the normalized mutual information I_N as the number of trials $s = n$ is increased. Physically, we ask how many different initial starting points in the energy landscape are required to achieve a certain desired threshold accuracy as measured by information theory measures.

In our multi-scale ("multiresolution") analysis, we solve $r = 100$ independent "replicas" (see Appendix A) and examine information theory correlations between the replicas and the planted ground state solutions. We schematically show such a set of independent solvers in Fig. 2 where stronger agreement among the replicas indicates a more robust solution. We compute the average inter-replica information correlations among the ensemble of replicas allowing us to infer a more detailed picture of the system beyond that of a single optimized solution. Specifically, *information theory extrema* as a function of T and γ (or other scale parameters in general) correspond to most relevant scale(s) of the system.

III. CONSTRUCTION OF EMBEDDED GRAPHS AND THE NOISE TEST

Similar to [48], we construct a "noise test" benchmark as a medium in which to study phase transitions in random graphs with embedded solutions [34, 35]. We define the system "noise" as intercommunity edges that connect a given node to communities other than its original or "best" community assignment. In general [34], it is not possible at the beginning of an attempted solution to ascertain which edges contribute to noise and which constitute edges within communities of the best partition(s).

For each benchmark graph, we divide N nodes into q communities with a power law distribution of community sizes $\{n_i\}$ given by n^β where $\beta = -1$. We then connect "intracommunity" edges at a high average edge density $p_{in} = 0.95$. Initially, the external edge density is zero,

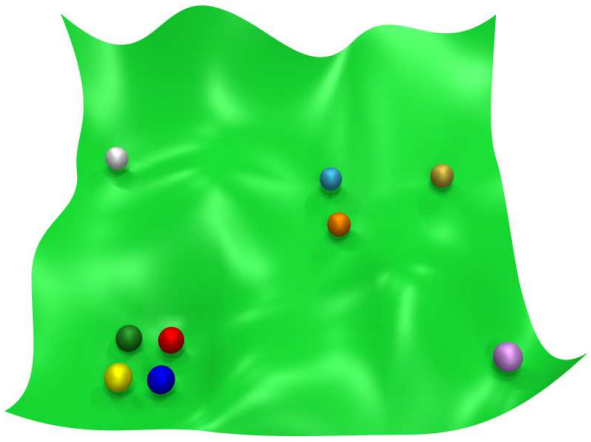


FIG. 2: (Color online) The schematic illustrates different minimizers attempting to solve a system. Colored spheres represent distinct minimizers ("replicas") that seek a (perhaps local) minimum of a cost function. In an easy system, multiple solution attempts will generally reach a good solution (such as the bottom left region of the landscape), but hard systems require more effort to solve accurately (that is, to achieve strong agreement between the replicas). Unsolvable regions restrict accurate solutions without extreme levels optimization (such as exhaustive search).

$p_{out} = 0$, so that we have perfectly decoupled clusters. To this system, we add random intercommunity edges at a density of $p_{out} < 0.5$. We define p_{in} (p_{out}) as the ratio of the number of intracommunity (intercommunity) edges over the maximum possible intracommunity (intercommunity) edges.

We define the *average external degree* of each node Z_{out} as the average number of links that a given node has with nodes in communities other than its own. Similarly, the *average internal degree* Z_{in} is defined as the average number of links to nodes in the *same* community, and $Z_{in} + Z_{out} = Z$ where Z is the average coordination number. Then we can explicitly write the internal and external edge densities

$$p_{in} = \frac{NZ_{in}}{\sum_{a=1}^{q} n_a(n_a - 1)}, \tag{2}$$

and

$$p_{out} = \frac{NZ_{out}}{\sum_{a=1}^{q}\sum_{b\neq a}^{q} n_a n_b}. \tag{3}$$

where n_a denotes the size of community a.

The communities in this construction are well defined, on average, at reasonable levels of noise ($p_{out} \lesssim 0.3$ depending on the typical community size n). As external links are progressively added to the system (p_{out} increases), the communities become increasingly difficult to detect. At some stage, enough noise is added and p_{out} is sufficiently high that the planted partition cannot be detected despite the fact that the optimal ground state is still well-defined. This transition often occurs sharply,

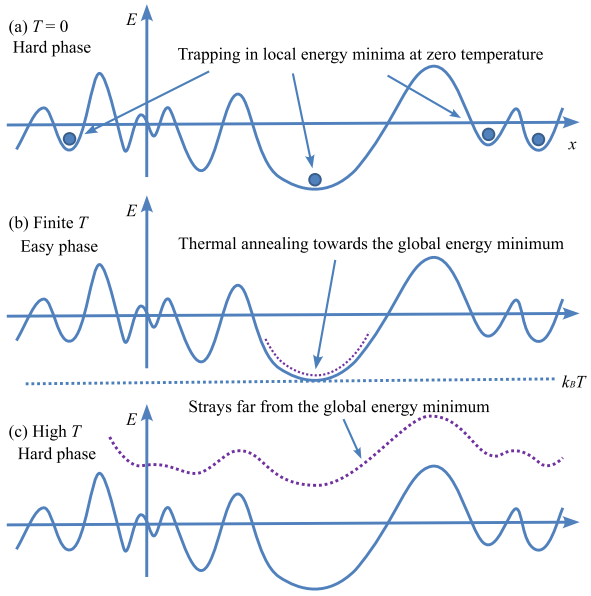


FIG. 3: (Color online) Panel (a) schematically illustrates in one dimension the easy and hard phases induced by the level of noise (extraneous intercommunity edges) encountered by a solver. Greedy algorithms are easily trapped in local energy minima above a certain noise threshold. We previously showed that the model of Eq. (1) is robust to noise [34] even with a greedy algorithm. Stochastic solvers such as a heat bath algorithm (see Appendix D) or simulated annealing enable one to circumvent the effects of some noise, but excessive levels will still thwart these solvers because meaningful partition information is obscured by the complexity of the energy landscape. Panels (b) and (c) schematically depict the easy and hard phases in terms of the temperature for the stochastic heat bath solver (see Appendix D). Above a graph-dependent threshold, the solver is less sensitive to local energy landscape features.

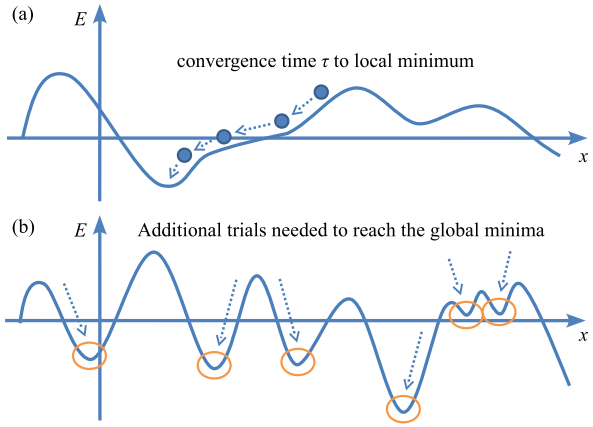


FIG. 4: (Color online) The figure schematically illustrates the convergence time of a solver in panel (a) and the effect of additional optimization trials in panel (b). Additional optimization trials are utilized in a "computational susceptibility" χ in order to numerically estimate the complexity of the energy landscape (see Appendix C).

particularly for large networks. We investigate the phase transition from the solvable to unsolvable phases at both low and high temperatures by means of the heat bath algorithm described in Appendix D.

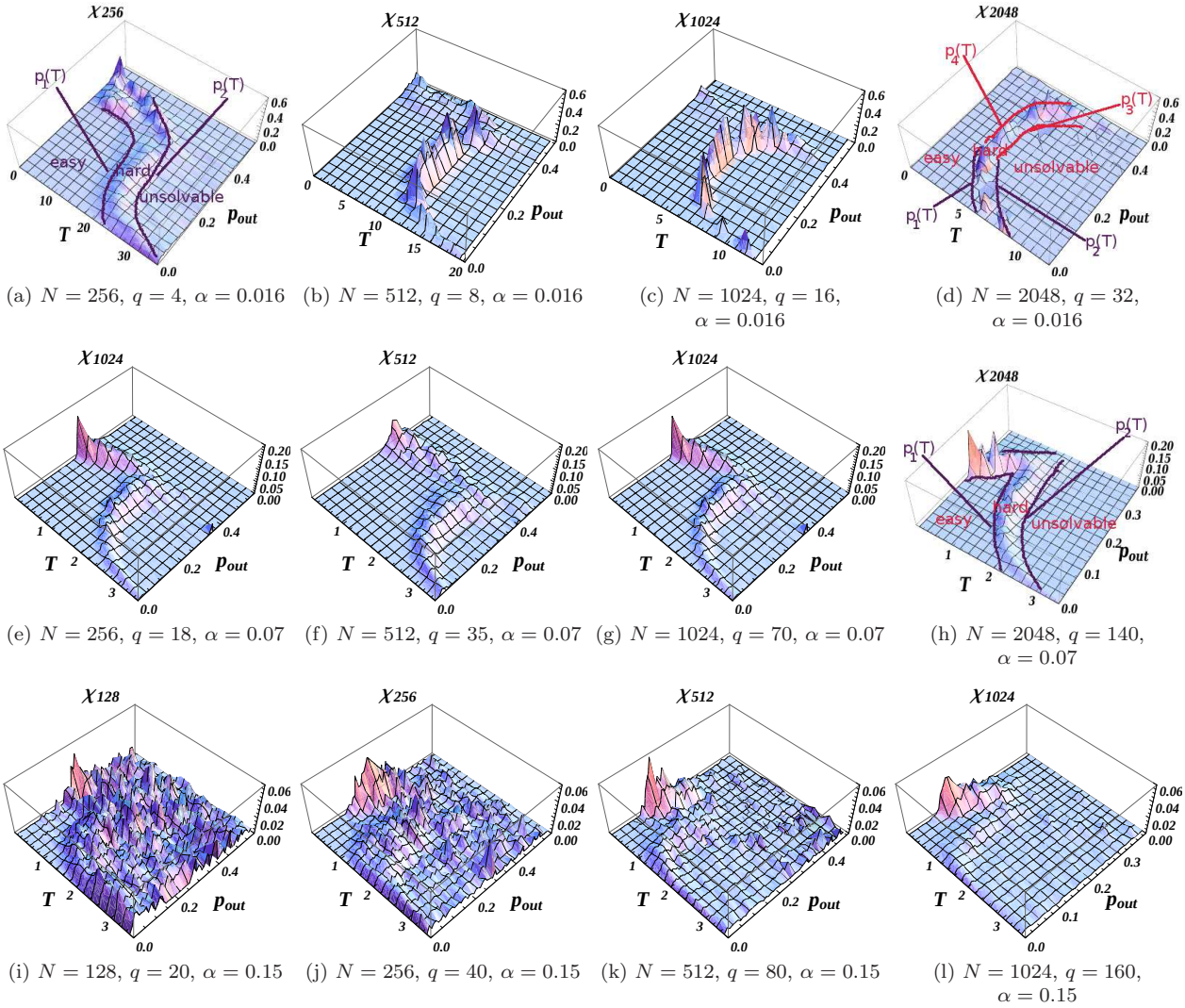
IV. SPIN GLASS TYPE TRANSITIONS

We previously reported [35] on the existence of *two* spin-glass-type transitions in the constructed graphs mentioned in Sec. III. Evidence for the transitions are observed in several measures such the accuracy of the solution obtained by means of the APM in Eq. (1) (and other models [9, 34] in general), the computational effort required to converge to a solution [11, 34], entropy effects, and others. Compared to another Potts-type qualtity function [9] utilizing a "null model" (a random graph used to evaluate the quality of a candidate partition), the APM exhibits a somewhat sharper transition as N is increased [34]. As alluded to above, two transitions are generally encountered as the noise value (or temperature) is increased. At fixed temperature T, as p_{out} is steadily increased from zero, the first onset of spin glass behavior first appears for values $p_1 \leq p_{out} \leq p_2$.

Figure 3(a) illustrates a one dimension characterization of the easy and hard phases in terms of the level of noise (extraneous intercommunity edges) encountered by a greedy solver. It is in this context that greedy algorithms are, in general, more easily trapped in local energy minima above a certain noise threshold. Stochastic solvers such a heat bath algorithm discussed in Appendix D or simulated annealing (SA) enable one to circumvent noise to some extent, but excessive levels will even thwart these more robust solvers because meaningful information is eventually obscured by the complexity of the energy landscape. Fig. 3(b,c) depict the easy and hard phases at low and high temperatures T, respectively, for our HBA (see Appendix D). Above a graph-dependent threshold, the solver is insensitive to local features, and it is unable to find an accurate solution.

We showed that Eq. (1) is robust to noise [34] leading to exceptional accuracy even with a greedy algorithm. Some other methods and cost-functions [7, 49] have also proven to be very accurate [18] with a greedy-oriented algorithm. While maximizing modularity [50] and a closely related cost function in [9] have proven to be accurate and productive, Refs. [36, 51, 52] have discussed problems associated with maximizing modularity in community detection. We briefly illustrated [34] a correspondence between the major transition experienced by Eq. (1) and a Potts model in [9]. We conjecture the existence of a related transition for random knots in Appendix G.

In Sec. IV A and IV B, we elaborate on the transitions using a *"computational susceptibility"* χ as defined in Appendix C(see also [11, 35]). In analogy with other physical susceptibility parameters, χ measures the response of the system to additional optimization effort. In the pres-



(a) $N = 256$, $q = 4$, $\alpha = 0.016$ (b) $N = 512$, $q = 8$, $\alpha = 0.016$ (c) $N = 1024$, $q = 16$, $\alpha = 0.016$ (d) $N = 2048$, $q = 32$, $\alpha = 0.016$

(e) $N = 256$, $q = 18$, $\alpha = 0.07$ (f) $N = 512$, $q = 35$, $\alpha = 0.07$ (g) $N = 1024$, $q = 70$, $\alpha = 0.07$ (h) $N = 2048$, $q = 140$, $\alpha = 0.07$

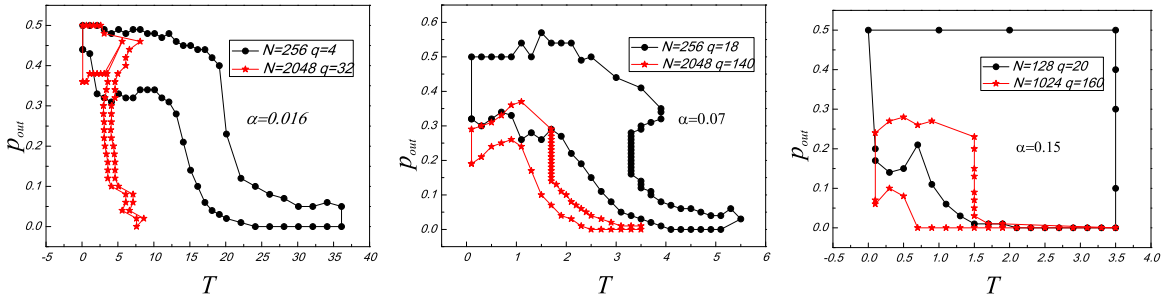
(i) $N = 128$, $q = 20$, $\alpha = 0.15$ (j) $N = 256$, $q = 40$, $\alpha = 0.15$ (k) $N = 512$, $q = 80$, $\alpha = 0.15$ (l) $N = 1024$, $q = 160$, $\alpha = 0.15$

FIG. 5: (Color online) Each panel shows a 3D plot $\chi(T, p_{out})$ as a function of temperature T and noise level p_{out} for systems with the indicated number of nodes N, communities q, and $\alpha = q/N$ ratio. In panels (a–h) for $\alpha = 0.016$ and 0.07, all plots show three clear phases, and the "ridges" at low and high temperatures mark the hard phase. The hard phase separates the easy phase (the flat region in the lower left corner with low temperature and low noise) from the unsolvable phase (the flat region in the upper right corner with high temperature and high noise). In panels (a)–(h), the ridges in $\chi(T, p_{out})$ become narrower as N increases. The area of the easy (hard) regions decreases (increases) from panel (a) to (d) and (e) to (h), respectively. In panels (a–d) for $\alpha = 0.016$, the hard phase at low temperature becomes less prominent from panel (a) to (d), but it becomes more prominent at high temperature. In panels (e–h) for $\alpha = 0.07$, the hard phase at low temperature becomes more prominent from panel (e) to (h), but it remains constant at high temperature. In panels (i–l) for $\alpha = 0.15$, only the larger systems with $N \geq 512$ show clear phases. The smaller systems with $N = 128$ in panel (i) and $N = 256$ in panel (j) show very noisy phases where only the easy phase can be readily determined, and the boundaries for the hard and unsolvable phases are difficult to pinpoint.

ence of multiple local minima [e.g., Fig. 4(b)], more trials may increase the accuracy. A higher χ indicates a more disordered, but navigable, energy landscape whereas a low χ indicates that additional optimization has less effect whether due to extreme disorder or a trivially solvable system. Finally in Sec. IV C, we illustrate the transitions using additional stability measures.

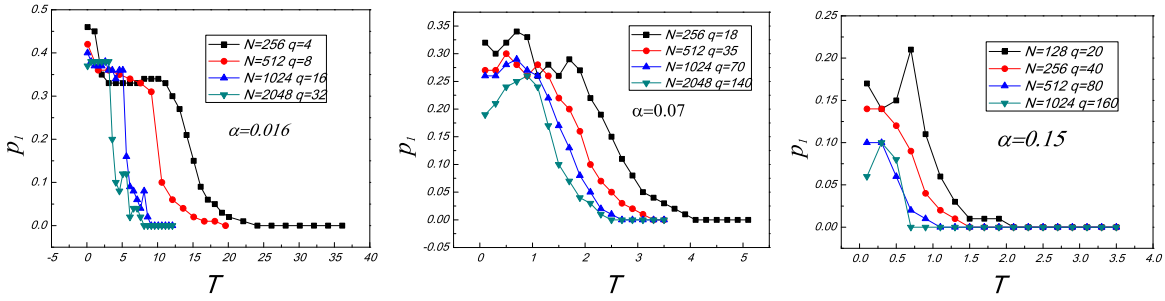
A. $\chi(T, p_{out})$ at fixed $\alpha = q/N$

We show the phase transitions in terms of three-dimensional (3D) plots with the computational susceptibility $\chi(T, p_{out})$ for a range of system sizes N and numbers of communities q. First, we fix the ratio $\alpha = q/N$ and study the phase transitions as N increases. Then we test a range of systems with fixed q as N increases.



(a) "hard" phase boundary for $\alpha = 0.016$ (b) "hard" phase boundary for $\alpha = 0.07$ (c) "hard" phase boundary for $\alpha = 0.15$

FIG. 6: (Color online) Corresponding to Fig. 5 and Sec. IV A, each plot depicts the boundaries of the hard phase for the system series with a fixed $\alpha = q/N$ ratio. Panels (a), (b), and (c) show the results for $\alpha = 0.016$, $\alpha = 0.07$, and $\alpha = 0.15$, respectively. System sizes range from $N = 256$ to 2048, and q varies from 4 to 160 as indicated in each plot. For each α, the area within hard phase boundary becomes progressively narrower indicating that the transitions from the easy to unsolvable phases are more clear in the thermodynamic limit.



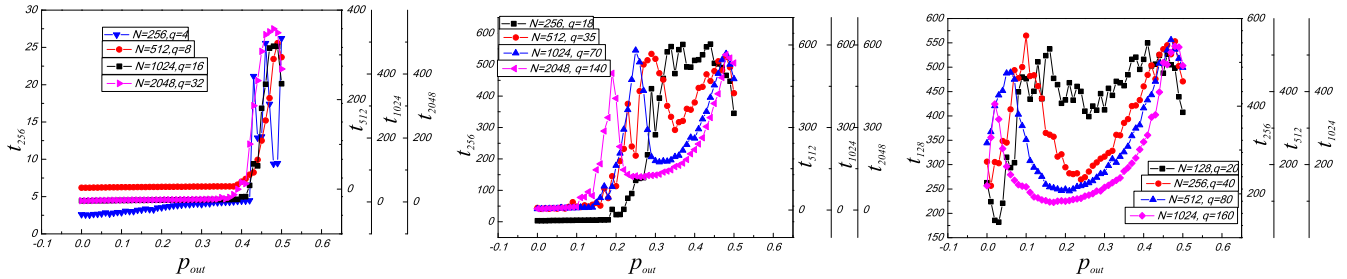
(a) $p_1(T)$ with $\alpha = 0.016$ (b) $p_1(T)$ with $\alpha = 0.07$ (c) $p_1(T)$ with $\alpha = 0.15$

FIG. 7: (Color online) Corresponding to Fig. 5 and Sec. IV A, each plot depicts the first phase transition point p_1 as a function of the temperature T for systems with a fixed ratio of $\alpha = q/N$. Panels (a), (b), and (c) show the results for $\alpha = 0.016$, $\alpha = 0.07$, and $\alpha = 0.15$, respectively. System sizes range from $N = 256$ to 2048, and q varies from 4 to 160 as indicated in each plot. All panels show that when α is fixed, the value of the first transition point p_1 decreases as the system size increases. This behavior further indicates that the system becomes more complex to solve in the thermodynamic limit.

1. $\chi(T, p_{out})$ at $\alpha = 0.016$

In Fig. 5 panels (a) through (d), we begin the analysis at a small $\alpha = q/N = 0.016$ ratio. The results for four system sizes are shown: $N = 256$, $N = 512$, $N = 1024$ and $N = 2048$ which maintain a fixed ratio of α across the respective rows. Each plot shows the "easy", "hard", and "unsolvable" phases.

The two "ridges" in each plot denote the hard phases. Generally, we may have two sets of spin-glass-type transitions into the hard phase as p_{out} or T is increased. For a given temperature T, hard phases generally appear for $p_1 \leq p_{out} \leq p_2$ (as marked in panels (a), (d), and (h) of Fig. 5). As we will re-iterate and allude to later, in some cases, these transitions correspond to an annealing effect wherein increasing the noise and/or temperature improves the accuracy of the solutions. Generally, higher temperature transitions may also appear [see, e.g., the V shaped ridge in panel (h) of Fig. 5 wherein p_1 and p_2 are non-monotonic in T]. As is similarly vividly seen in panel (d), the hard phase boundaries may become multi-valued and "bow" as a function of p_{out}; such a situation corresponds to a re-entrant transition whence the system goes from the unsolvable disordered phase via a hard phase extending from $p_1(T) < p_{out} < p_2(T)$ to the ordered easy phase and then, via a second higher noise hard phase ($p_3(T) < p_{out} < p_4(T)$) back to the unsolvable disordered phase as p_{out} is increased at a fixed temperature T. Later on, we will corroborate the location of these phase boundaries by the use of other metrics including disparate information theoretic measures and thermodynamic quantities such as the entropy and energy. For a fixed value of the ratio of q/N, the height of the first ridge at low temperature decreases as the system size N increases while the height of the second ridge at high temperature increases in the same process. This finite size scaling behavior for the hard phase at high temperature indicates that the phase transition at high temperature



(a) τ with $\alpha = 0.016$ (b) τ with $\alpha = 0.07$ (c) τ with $\alpha = 0.15$

FIG. 8: (Color online) Corresponding to Fig. 5 and Sec. IV A, the convergence time τ [see Fig. 4(a)] as a function of noise p_{out} at zero temperature for the systems with a fixed ratio of $\alpha = q/N$. Panels (a), (b), and (c) show the results for $\alpha = 0.016$, $\alpha = 0.07$ and $\alpha = 0.15$, respectively. System sizes range from $N = 256$ to $N = 2048$, and q varies from 4 to 160 as indicates in each plot. The noise level p_{out} at the first peak of the convergence time corresponds to the first transition point p_1 in Fig. 7 at zero temperature. As the system size increases, the first peak in the convergence time moves to the left. They share the same trend as in Fig. 6 and Fig. 7.

exists in the thermodynamic limit. However, the phase transition at low temperature will disappear in the same limit. In the meantime, the ridge in the high temperature will gradually expand into the low temperature region as the system size increases. Thus, for the systems with the small ratio of α, the phase transition will exist in almost the entire temperature range in the thermodynamic limit (see Sec. V).

The "easy" phase shrinks and the unsolvable phase expands as q increases for fixed $\alpha = \frac{q}{N}$. The approximate area of the easy phase on the left corner in panel (a) is in the range of $T \in (0, 20)$ and $p_{out} \in (0, 0.4)$. The area of the unsolvable phase on the right upper corner is in the range of $T \in (20, +\infty)$ and $p_{out} \in (0, 0.4)$. As the system size increases from $q = 4$ in panel (a) to $q = 16$ in panel (c) for a fixed $\alpha = \frac{q}{N}$, the area of the easy phase shrinks to the range of $T \in (0, 5)$ and $p_{out} \in (0, 0.4)$ while the unsolvable phase expands to $T \in (5, +\infty)$ and $p_{out} \in (0, 0.4)$. As the number of communities further increases to $N = 32$ in panel (d), the easy phase further shrinks to the range of $T \in (0, 4)$ and $p_{out} \in (0, 0.4)$ while the unsolvable phase expands to $T \in (4, +\infty)$ and $p_{out} \in (0, 0.4)$. We note that the range of p_{out} for the easy phase does not decrease as the number of communities increases.

Low temperature transitions correspond to an "order by disorder" type effect where fluctuations render the system ordered or solvable. As the temperature T increases, the system may veer towards its global minimum by annealing. In such cases, by virtue of entropic fluctuations, quenching is thwarted and the system may probe low lying states and achieve an ordered configuration.

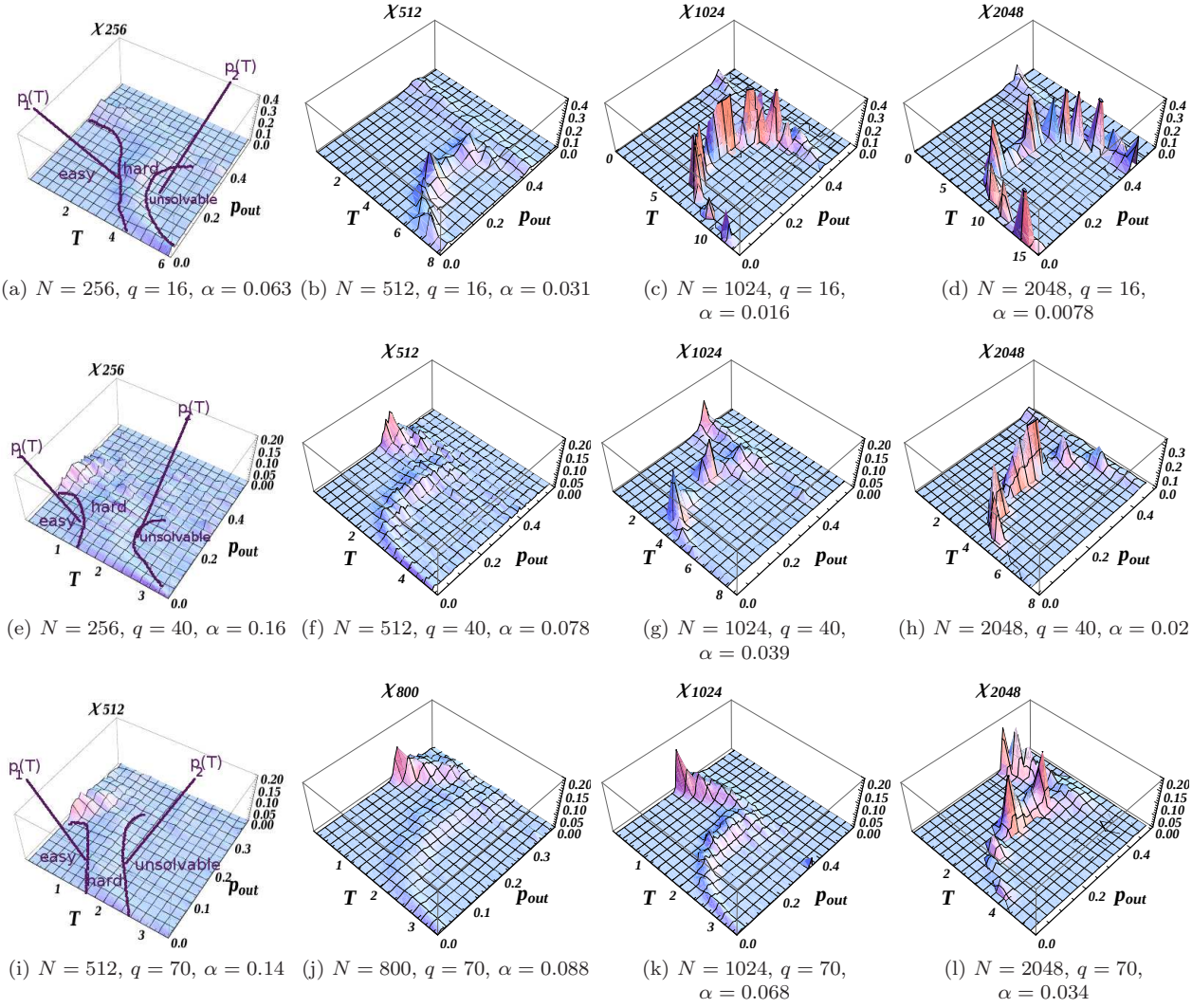
In order to track the range of the hard phases, we further display a set of "boundary" plots in Fig. 6 as well as the first transition point p_1 as the function of temperature in Fig. 7. For the system series with the fixed $\alpha = 0.016$ discussed above, the 2D "hard phase" boundaries and the values of the first transition points are in panel (a) of Fig. 6 and Fig. 7, respectively.

In Fig. 6(a), the area of the hard phase shrinks, and its area at high temperature becomes narrower as the system size increases. Specifically, the width of the hard phase for $N = 256$ is about $\Delta T = 6$, while it only extends to $\Delta T = 1$ for the $N = 2048$. Together with the 3D phase diagrams in panels (a)–(d) of Fig. 5, we conclude that the hard phase at the high temperature becomes sharper in the thermodynamic limit.

The boundaries of the hard phase at low temperature are more easily seen in Fig. 7(a) where we plot the first transition point p_1 as the function of temperature T for a range of systems. The plots confirm the observations in Fig. 5(a)–(d) regarding the constant p_{out} range. That is, the range of p_{out} for the easy phase does not decrease as the system size increases [in Fig. 7(a), p_1 collapses before $T \leq 5$ for all the systems]. This behavior hints that the first transition point p_1 at low temperature and small α remains constant in the thermodynamic limit.

As depicted in Fig. 4(a), the convergence time τ provides another view of the phase transition. We plot τ as a function of noise level p_{out} in Fig. 8(a) for systems with a fixed ratio of $\alpha = q/N = 0.016$. The value p_{out} at the first peak of the convergence time in each system is consistent with the first transition point p_1 observed in Fig. 7(a). As the number of communities q increases, the peak convergence time shifts to the left, which corresponds to the lower value of p_1.

For a given noise level p_{out}, the highest temperature at which the system is still solvable tends to zero as the number of communities q increases. As we will demonstrate in this work for myriad examples (see also [59]), the critical temperature beyond which the system is unsolvable is $T^* \approx 1/\log q$. As the number of communities $q \to \infty$, the entire system becomes unsolvable.



(a) $N = 256$, $q = 16$, $\alpha = 0.063$ (b) $N = 512$, $q = 16$, $\alpha = 0.031$ (c) $N = 1024$, $q = 16$, $\alpha = 0.016$ (d) $N = 2048$, $q = 16$, $\alpha = 0.0078$

(e) $N = 256$, $q = 40$, $\alpha = 0.16$ (f) $N = 512$, $q = 40$, $\alpha = 0.078$ (g) $N = 1024$, $q = 40$, $\alpha = 0.039$ (h) $N = 2048$, $q = 40$, $\alpha = 0.02$

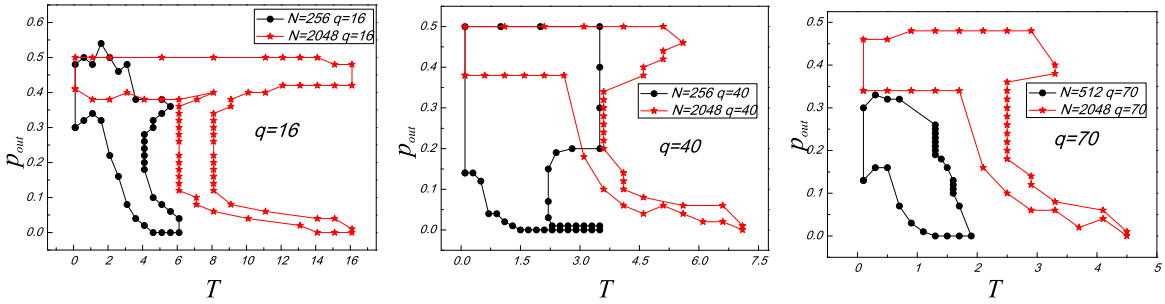
(i) $N = 512$, $q = 70$, $\alpha = 0.14$ (j) $N = 800$, $q = 70$, $\alpha = 0.088$ (k) $N = 1024$, $q = 70$, $\alpha = 0.068$ (l) $N = 2048$, $q = 70$, $\alpha = 0.034$

FIG. 9: (Color online) Similar to Fig. 5, we plot of $\chi(T, p_{out})$ as a function of temperature T and noise level p_{out} for systems with the indicated number of nodes N, communities q, and $\alpha = q/N$ ratio. Here, q is fixed for each row series, and we vary α (rows) to examine the behavior as N increases (columns). The heights of the susceptibility peaks at higher T increase across each series as N increases whereas the heights at low T are relatively constant. The $N = 256$ node systems do not show clear hard or unsolvable phases, but the transitions are strong at high temperature for most panels in the second and third columns of plots.

2. $\chi(T, p_{out})$ at $\alpha = 0.07$

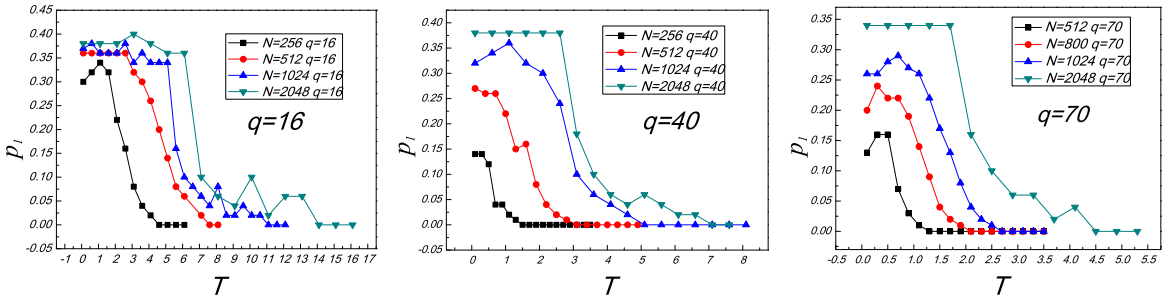
For $\alpha = 0.07$, the phase transitions are presented in Fig. 5 panels (e) through (h). The phases in panel (e) are noisy compared to panels (f) through (h), and all of the systems are more complicated than the plots with $\alpha = 0.016$. As N increases, the phase transitions become more clear. However, contrary to panels (a) through (d), the phase transition at low temperature becomes more prominent as N increases, and the transition at high temperature stays roughly constant. Specifically, the height of the susceptibility peak at low temperature increases from $\chi = 0.01$ at $N = 256$ in panel (e), $\chi = 0.05$ at $N = 512$ in panel (f), $\chi = 0.1$ for $N = 1024$ in panel (g), and finally reaches $\chi = 0.2$ in panel (h) with $N = 2048$. The phase transitions in this series appear to be persistent.

The easy phase (lower left of each panel) decreases in area as the system size increases. This is the same trend that was observed in the previous $\alpha = 0.016$ series implying that the easy phase will tend to decrease in the thermodynamic limit up to a threshold (see Sec. V). Specifically, the easy phase in the smallest system in panel (e) covers the range of $T \in (0, 3)$ and $p_{out} \in (0, 0.3)$ while in the large system in panel (h) covers $T \in (0, 1.5)$ and $p_{out} \in (0, 0.2)$. The range for p_{out} in the easy phase decreases as the N increases which differs from the $\alpha = 0.016$ data where the noise p_{out} stayed at a



(a) hard phase boundary for $q = 16$ (b) hard phase boundary for $q = 40$ (c) hard phase boundary for $q = 70$

FIG. 10: (Color online) Corresponding to Fig. 9 and Sec. IV B, each plot depicts the boundaries of the hard phase for the system series with a fixed number of communities q where panels (a), (b), and (c) correspond to $q = 16$, 40, and 70, respectively. System sizes range from $N = 256$ to 2048 as indicated. For each q, the area of the hard phase becomes progressively narrower which indicates clearer transitions from the easy to unsolvable phases in the thermodynamic limit.



(a) $p_1(T)$ with $q = 16$ (b) $p_1(T)$ with $q = 40$ (c) $p_1(T)$ with $q = 70$

FIG. 11: (Color online) Corresponding to Fig. 9 and Sec. IV B, each plot depicts the first phase transition point p_1 as a function of temperature T for systems with a fixed q. Panels (a), (b), and (c) show the results for $q = 16$, 40, and 70, respectively. System sizes range from $N = 256$ to 2048 as indicated in each plot. All panels show that the first transition point increases as the system size increases which is consistent with the complexity trend of the system series.

roughly constant range of $p_{out} \in (0, 0.4)$. In both series for $\alpha = 0.016$ and 0.07, the value of the initial transition point p_1 decreases in the thermodynamic limit.

The corresponding 2D plots of the hard phase boundaries and the first transition points p_1 are displayed in Fig. 6(b) and Fig. 7(b), respectively. For the series with $\alpha = 0.07$ in Fig. 6(b), the area of the hard phase becomes narrower at both low and high temperatures as the system size increases. In detail, the width of the hard phase for $N = 256$ is about $\Delta T = 1.3$, while the width shrinks to about $\Delta T = 0.3$ at $N = 2048$. Together with the 3D phase diagrams in Fig. 5(e)–(h), the phase transitions become sharper in the thermodynamic limit.

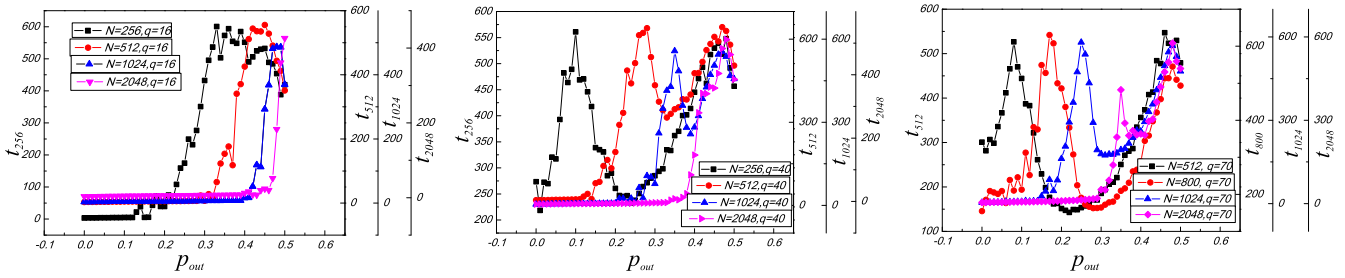
As shown in Fig. 7(b), the first transition point p_1 decreases as the system size increases, even in the low temperature limit. This is consistent with the first peak of the convergence time τ at zero temperature in Fig. 8(b). This indicates that the system becomes progressively harder to solve in the thermodynamic limit over the whole temperature range.

3. $\chi(T, p_{out})$ at $\alpha = 0.15$

In panels (i) through (l) of Fig. 5, $\alpha = 0.15$ and the clusters are smaller on average resulting in systems that are more difficult to solve. In panels (i) and (j), almost the entire region is covered by small peaks which indicates mixing of the hard and unsolvable phases thus making the phase boundaries hard to detect.

The flat easy regions are recognizable in all panels, but the area is small relative to the previous cases and becomes even smaller as N increases into panel (l). In panel (i), the flat easy region is roughly triangular with legs along $T \in (0, 1.5)$ and $p_{out} \in (0, 0.2)$. The easy region shrinks to a smaller triangle along $T \in (0, 0.2)$ and $p_{out} \in (0, 0.2)$ in panel (j) and (k). In panel (l), it further shrinks to $T \in (0, 1)$ and $p_{out} \in (0, 0.1)$. The easy phase shrinks for both p_{out} and T as N increases which further indicates that the initial transition point p_1 decreases substantially in the thermodynamic limit.

The corresponding plots of the hard phase boundaries and the first transition points p_1 are displayed in Figs.



(a) τ with $q = 16$ (b) τ with $q = 40$ (c) τ with $q = 70$

FIG. 12: (Color online) Corresponding to Fig. 9 and Sec. IV B, the convergence time τ [Fig. 4(a)] as a function of noise p_{out} for systems with fixed q. Panels (a), (b) and (c) show the results for $q = 16$, 40, and 70, respectively. System sizes vary from $N = 256$ to 2048 as indicated in each plot. The noise level p_{out} at the first peak of the convergence time corresponds to the initial transition point p_1 in Fig. 11 at zero temperature. As the system size increases, the first peak in the convergence time moves to the right. They share the same trend as in Figs. 10 and 11.

6(c) and 7(c), respectively. From Fig. 6(c), the area of the hard phase shrinks in the thermodynamic limit. The hard phase is more identifiable relative to the unsolvable region as N increases. The initial transition point p_1 drops as N increases as shown in Fig. 7(c). The convergence time τ for the systems with the fixed ratio of $\alpha = q/N = 0.15$ at zero temperature is shown in Fig. 8(c) where the first peak of τ shifts to the left as the system size increases. This is consistent with the trend observed in Fig. 7(c). We further show in Fig. 14 and Fig. 15 that the first transition points in "computational susceptibility", energy, entropy, convergence time and normalized mutual information are consistent with each other.

In Fig. 16, we provide plots of scaled waiting correlation function data which clearly indicate *spin glass type collapse*. The collapse is best at the center of the computational susceptibility ridge Fig. 16(b). The collapse persists up to the ends of the susceptibility ridge (e.g., $p_{out} = p_1$ in Fig. 16(a)) and is no longer valid outside the susceptibility ridge (e.g., $p_{out} = 0.26 > p_2 = 0.24$ in Fig. 16(c)).

B. $\chi(T, p_{out})$ at fixed q

We fix the number of communities at $q = 16$, 40, or 70 and increase the system size N from 256 to 2048. The plots of computational susceptibility $\chi(T, p_{out})$ for $q = 16$ series of systems are shown in panels (a) through (d) of Fig. 9. As in Sec. IV A, the ridges indicate hard phases which become more prominent as N increases while the ridges at low temperature remain at relatively low constant values.

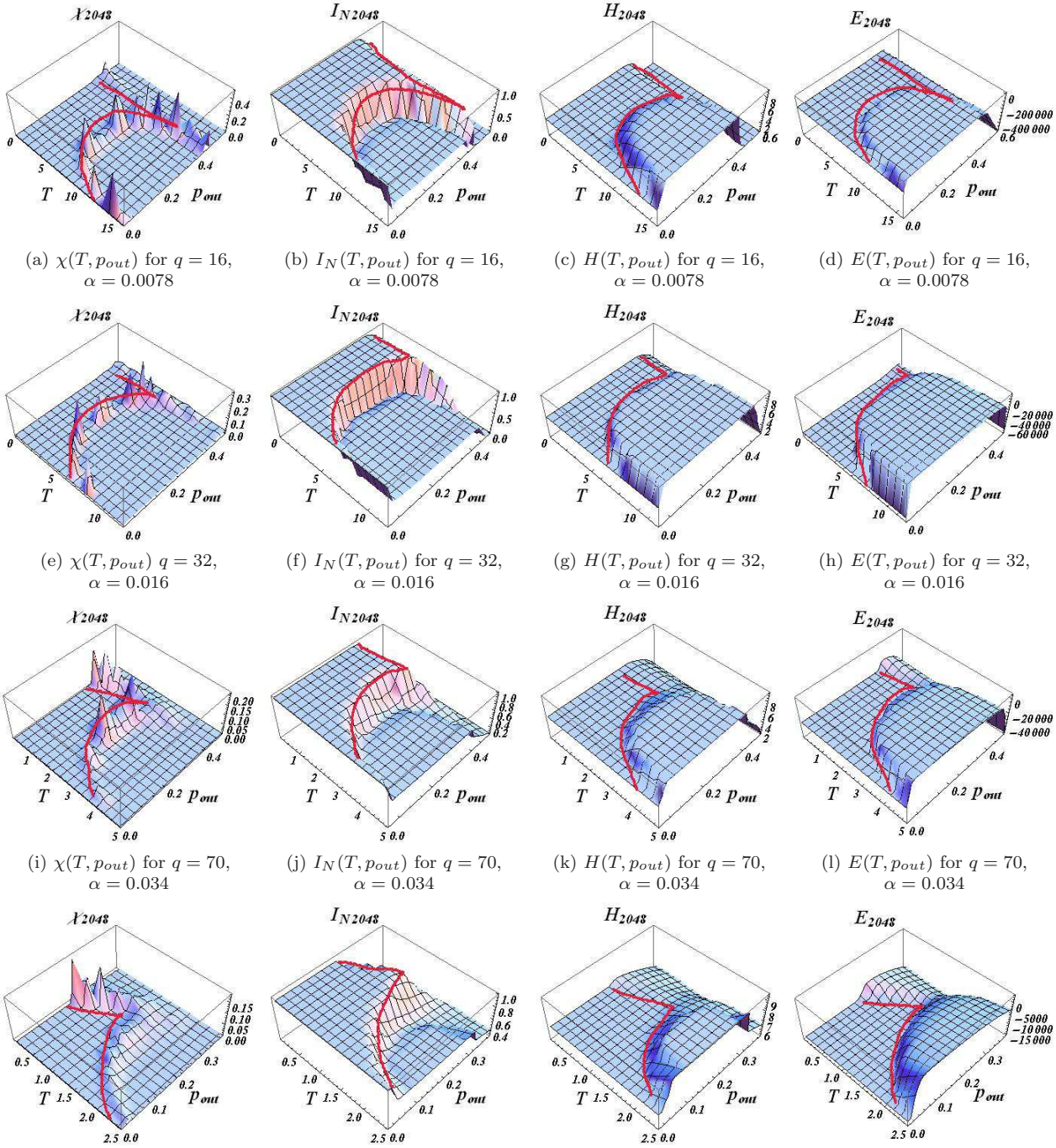
The areas of the easy phases on the lower left corner expand as the system size increases from panel (a) to (d). This trend of increasing area is the reverse of the behavior in the fixed α systems systems in Sec. IV A. This is easy to understand since, q increases with N here, and the high internal edge density p_{in} causes the larger clusters to be more strongly defined.

We increase the number of communities to $q = 40$ for the systems in panels (e) through (h). N varies from 256 to 2048, and $\alpha = q/N$ decreases as N increases so that the systems again become less complicated because the communities become more strongly defined. The hard and unsolvable phases in the small $N = 256$ system in panel (e) are difficult to distinguish. Only the easy phase can be easily identified by noting the flat region on the lower left of each panel. The "computational susceptibility" $\chi(T, p_{out})$ peaks at increasing heights at both the low and high temperatures from panels (f) to (h) indicating that the phase transitions become more prominent as the system size increases.

We further increase the number of communities to $q = 70$ and study the phase transitions for the same range of system sizes. The hard phase at high temperature in panel (i) is difficult to detect. $\chi(T, p_{out})$ clearly shows the three phases in panels (j) and (k). The easy phases again become larger as the system size increases. $\chi(T, p_{out})$ in the hard phase increases as N increases indicating that the phase transitions at both low and high temperatures are more obvious from panel (i) to (k).

In Figs. 10 and 11, we also show corresponding 2D plots for the boundaries of the hard phase and the first transition point p_1 as the function of temperature T. In Fig. 10, the area of the hard phase becomes narrower as the system size increases. At $q = 40$, for example, the width of the hard phase for the smallest system at $N = 256$ is about $\Delta T = 1.5$. As N increases, the hard phase width shrinks to $\Delta T = 1$ at $N = 512$ and down to $\Delta T = 0.5$ for $N = 2048$ which further indicates that the phase transition becomes sharper in the thermodynamic limit. In Fig. 11, the first transition point p_1 increases over the entire temperature range as N increases. This behavior is consistent with the system complexity trend as previously mentioned.

In Fig. 12, we further plot the convergence time τ as the function of noise p_{out} for a fixed number of commu-



(a) $\chi(T, p_{out})$ for $q = 16$, $\alpha = 0.0078$ (b) $I_N(T, p_{out})$ for $q = 16$, $\alpha = 0.0078$ (c) $H(T, p_{out})$ for $q = 16$, $\alpha = 0.0078$ (d) $E(T, p_{out})$ for $q = 16$, $\alpha = 0.0078$

(e) $\chi(T, p_{out})$ $q = 32$, $\alpha = 0.016$ (f) $I_N(T, p_{out})$ for $q = 32$, $\alpha = 0.016$ (g) $H(T, p_{out})$ for $q = 32$, $\alpha = 0.016$ (h) $E(T, p_{out})$ for $q = 32$, $\alpha = 0.016$

(i) $\chi(T, p_{out})$ for $q = 70$, $\alpha = 0.034$ (j) $I_N(T, p_{out})$ for $q = 70$, $\alpha = 0.034$ (k) $H(T, p_{out})$ for $q = 70$, $\alpha = 0.034$ (l) $E(T, p_{out})$ for $q = 70$, $\alpha = 0.034$

(m) $\chi(T, p_{out})$ for $q = 140$, $\alpha = 0.068$ (n) $I_N(T, p_{out})$ for $q = 140$, $\alpha = 0.068$ (o) $H(T, p_{out})$ for $q = 140$, $\alpha = 0.068$ (p) $E(T, p_{out})$ for $q = 140$, $\alpha = 0.068$

FIG. 13: (Color online) Plots of the computational susceptibility χ (column one), NMI I_N (column two), Shannon entropy H (column three), and energy E (column four) as functions of temperature T and intercommunity noise p_{out}. System sizes all use $N = 2048$, and q varies from 16 to 140 in different rows. All plots show the easy, hard, and unsolvable phases often by rapid shifts in the respective measures. The marked (red, color online) line highlights the onset of the hard phase at $p_{out} = p_1(T)$ in each measured quantity where we note that the boundaries match well across each row.

nities q at zero temperature. p_{out} for the first peak of the convergence time matches the first transition point p_1 in Fig. 11. As the system size increases, the peak moves to the right. This is also consistent with Fig. 11 where the system becomes less complicated as N increases.

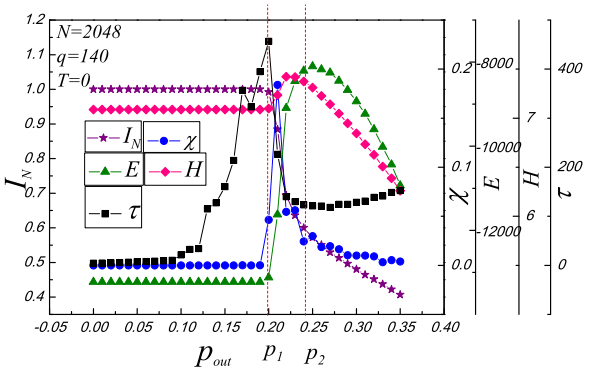


FIG. 14: (Color online) The plots of susceptibility χ, convergence time τ, energy E, accuracy I_N and the Shannon entropy H in terms of noise p_{out} for the system $N = 2048$ and $q = 140$ at a zero temperature. All the plots show three phases as noise varies: (1) Below $p_1 = 0.2$, the system can be solved in this "easy" region (e.g., the accuracy is $I_N = 1$); (2) When $0.2 < p_{out} < 0.24$, where the benefit of extra trials is the largest, it's "hard" to solve the system without misplacing nodes (e.g., χ, E and H achieve the peak) ; (3) Above $p_2 = 0.24$, the system is "impossible" to be perfectly solved. $[p_1, p_2]$ are generous bounds in transition crossover regions. Note that the two transitions are demonstrated to be of spin-glass-type by observing the scaling of the correlation function between $[p_1, p_2]$ in Fig. 16.

C. Other information theoretic and thermodynamic quantities

We further fortify and provide our results of the phase diagram of our systems as ascertained via other information theoretic and thermodynamics quantities. These measures include the average normalized mutual information I_N between replica pairs, Shannon entropy H, and energy E as shown in Fig. 13. We additionally show the corresponding computational susceptibility χ from Fig. 5 or 9 for comparison. All panels are for a system of size $N = 2048$. In panels (a) through (d), $q = 16$ which corresponds to Fig. 5(d). Panels (e) through (h) plot results for $q = 32$ with $\alpha = 0.015$ which corresponds to Fig. 9(d). Panels (i) through (l), display the results for $q = 70$ which corresponds to Fig. 5(l). Finally, Panels (m) through (p) display results for $q = 140$ and $\alpha = 0.07$ corresponding to Fig. 9(h).

All panels consistently display the three different complexity phases: the "easy" (flat region, lower left), "hard" (varied central regions), and "unsolvable" phases (far right or top). The existence of the hard phase is reflected by the ridges at both low and high temperatures in the susceptibility χ plot which often corresponds rapids shifts (up or down) in the other measures. In each plot, the red line serves as a guide to the eye to emphasize the boundaries between different phases. The phase boundaries adduced from the different measures agree with each other.
.

In Ref. [35], we also demonstrated the spin glass character of the phase transition by observing the exceptional collapse of time autocorrelation curves (over four orders of magnitude of time at high and low temperatures) in the vicinity of the hard phase. We further elucidated on evidence regarding phase transitions [35] in identifying community structure via a dynamical approach (some other dynamical methods include [13, 45]) where "chaotic-type" transitions that we speculated upon may extend into the node dynamics for large systems.

V. NON-INTERACTING CLIQUES

As depicted in Fig. 17, we analytically estimate a minimum transition temperature by examining a system with q non-interacting cliques. In panel (a), each of the q communities consists of l nodes which are maximally connected, but no noise exists between these cliques. The presence of noise will, in general, lower the temperature $T_\times$ of the transition point which manifests as departure from the easy phase in certain regions of Figs. 5 and 9.

Within our algorithm and model, communities do not interact in an explicit sense. In addition, with this model problem the situation is greatly simplified because no edges are assigned between cliques, so we use Eq. (1) to calculate the partition function of the system by counting the energy contribution of all edges within each cluster over the number of combinations for partitioning the clusters. As a further simplification, we also set the energy contribution for a single edge to be -2 so that the Hamiltonian gives an energy of -1 for each edge.

A. Partition function

First, we investigate the smallest non-trivial clique size with $l = 3$ nodes. The partition function for the decoupled cliques is,

$$\mathbf{Z} = (\mathbf{Z}_l)^q = \sum_{\sigma_i,\sigma_j} e^{-\beta H_{i,j}} \tag{4}$$

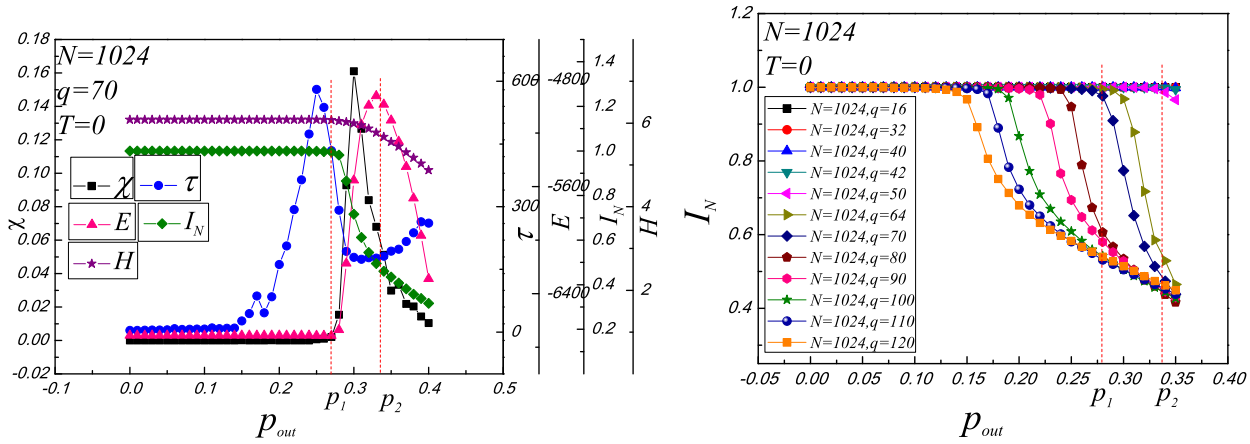
where $\mathbf{Z}_l$ is the partition function for a single clique and $\beta = 1/T$ is the inverse temperature. Considering the $l = 3$ cluster combinations depicted in Fig. 17(b), $\mathbf{Z}_3$ is

$$\mathbf{Z}_3 \;=\; qe^{6\beta} + 3q(q-1)e^{2\beta} + q(q-1)(q-2). \tag{5}$$

The first term represents the optimal local cluster solution, and the sum of the remaining terms accounts for the remaining sub-optimal local partitions. We define ω_l as the ratio of Boltzmann weights of the sub-optimal partitions to the optimal solution. For $\mathbf{Z}_3$, the ratio ω_3 is

$$\omega_3 = \frac{q(q-1)\left[3e^{2\beta} + (q-2)\right]}{qe^{6\beta}}. \tag{6}$$

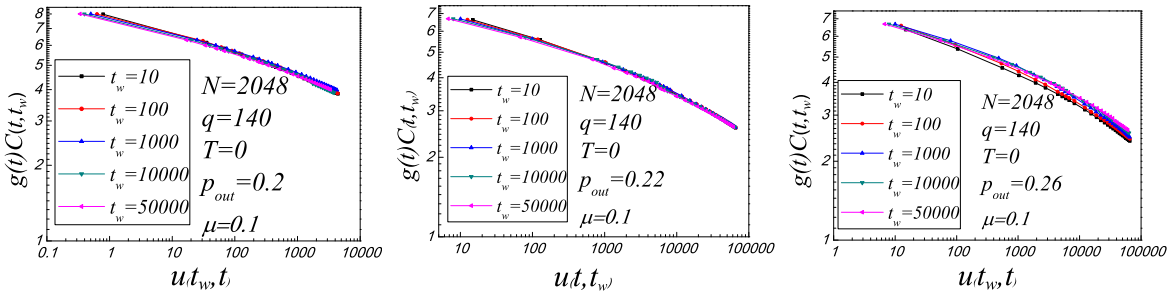
$\omega_l < 1$ indicates that the optimal solution is dominant, while $\omega_l \to \infty$ means the system is disordered. We can



(a) χ, τ, E, H, I_N for the system of size $N = 1024$

(b) I_N for the system of $N = 1024$ with different q

FIG. 15: (Color online) (a), The plots of susceptibility χ, convergence time τ, energy E, accuracy I_N, Shannon entropy H in terms of noise p_{out} for the system $N = 1024$ and $q = 70$ at a zero temperature. (b), The normalized mutual information I_N in terms of noise p_{out} for a series of systems with the size of $N = 1024$ but different number of communities q. From both plots, we are able to detect the first and second transition point p_1 and p_2. p_1 is the point where the I_N drops from 1, χ increases from 0, τ achieves the peak, E and H increases from some constant value. p_2 is the position where the I_N curves with different number of communities collapse shown in (b). p_2 also corresponds to the peak of energy and entropy as shown in (a).



(a) $p_{out} = 0.2$ is within the zero temperature "hard" phase, where the collapse is perfect.

(b) $p_{out} = 0.22$ is within the zero temperature "hard" phase, where the collapse is perfect.

(c) $p_{out} = 0.26$ is within the zero temperature "unsolvable" phase, where the collapse is poor.

FIG. 16: (Color online). We show a collapse of the correlation curves for different waiting times t_w for a system with $N = 2048$ nodes, $q = 140$ communities. p_{out} varies from 0.2 in panel (a) to 0.26 in panel (c). The first and second transition points for this system are $p_1 = 0.2$ and $p_2 = 0.24$. The temperature is $T = 0$. The vertical axis is $g(t)C(t_w, t)$ where $g(t) = 8 - \log_{10}(t)$, $C(t_w, t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\sigma_i(t_w),\sigma_i(t_w+t)}$ is the correlation function. The horizontal axis is $u(t_w, t) = \frac{1}{1-\mu}[(t + t_w)^{1-\mu} - t_w^{1-\mu}]$ where $\mu = 0.1$. The noise $p_{out} = 0.2$ (a) and $p_{out} = 0.22$ (b) lie within the "hard" region where the collapse of correlation function is perfect. The noise $p_{out} = 0.26$ (c) is above the second transition point p_2 in the "unsolvable" region, where the collapse becomes poor. That the collapse of the correlation function starts to degrade right after the second transition point p_2 at zero temperature indicates that this transition is of the spin-glass type.

define $\omega_l = 1$ as the transition point from the ordered phase to the disordered phase, and the corresponding "crossover" temperature $T_\times$ is found by solving the transcendental equation

$$3(q-1)e^{-4/T_\times} + (q-1)(q-2)e^{-6/T_\times} = 1. \quad (7)$$

In the limit of large q, this equation simplifies to

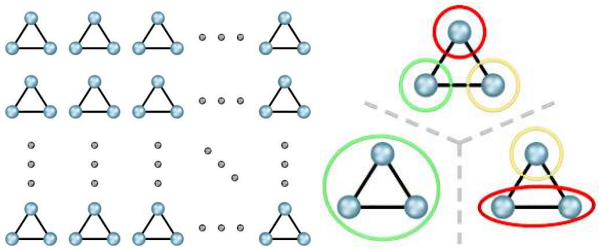
$$q^2 e^{-6/T_\times} \simeq 1, \quad (8)$$

which yields our estimate for the crossover temperature

$$T_\times \simeq \frac{3}{\log q} \quad (9)$$

for the $l = 3$ clique system. This is in agreement to general trend on accuracy in [59].

If we generalize to arbitrary clique size l, the corre-



(a) (b)

FIG. 17: (Color online) Panel (a) depicts q independent cliques (maximally connected clusters). Panel (b) indicates the different combinations of $l = 3$ nodes which must be summed (including three copies of the 2-1 configuration) in order to determine the partition function for a single clique.

sponding partition function for a single clique becomes

$$\begin{aligned} \mathbf{Z}_l &= q e^{2\beta\binom{l}{2}} + lq(q-1)e^{2\beta\binom{l-1}{2}} \\ &+ \frac{l(l-1)}{2} q(q-1)(q-2)e^{2\beta\binom{l-2}{2}} \\ &+ \cdots + q(q-1)(q-2)\cdots(q-l). \end{aligned} \tag{10}$$

Again, the first term in Eq. (10) is the Boltzmann weight of the optimal clique partition, and the other terms sum the weights of the incorrect partitions. ω_l is

$$\omega_l = \frac{\sum_{k=1}^{l} \binom{l}{k}\binom{q}{k+1}(k+1)! e^{2\beta\binom{l-k}{2}}}{q e^{2\beta\binom{l}{2}}}, \tag{11}$$

and $\omega_l = 1$ returns the cross-over temperature $T_\times$ for arbitrary cliques of size l. We summarize the crossover temperature relations in column one of Table I where we express $e^{2/T_\times}$ in terms of powers of q for several values of l. The general relation is

$$T_\times \simeq \frac{l}{\log q}. \tag{12}$$

B. Symmetry Breaking

We can inquire about the crossover temperature $T_\times$ from another perspective. Take two nodes i and j in the same clique. If the probability that a solution assigns them to the same community is high, then the system is in the "ordered" state. If this probability is $1/q$, the system is in its "disordered" phase. We can define a crossover temperature $T_\times^{(1/q)}$ at which the probability of node i and j being in the same cluster exceeds $1/q$ and thus symmetry between Potts spins is broken. This probability $P(\sigma_i = \sigma_j) = \langle \delta_{\sigma_i,\sigma_j} \rangle$ is

$$P(\sigma_i = \sigma_j) = \frac{\mathrm{Tr}_{\sigma_i} \delta_{\sigma_i,\sigma_j} e^{-\beta H}}{\mathrm{Tr}_{\sigma} e^{-\beta H}}, \tag{13}$$

l	$e^{2/T_\times}$	$e^{2/T_\times^{(1/q)}}$	$e^{2/T_\times^{(p)}}$
2	q	q	pq^2
3	$q^{2/3}$	$q^{2/3}$	$p^{1/3}q$
4	$q^{1/2}$	$q^{1/2}$	$p^{1/6}q^{2/3}$
5	$q^{2/5}$	$q^{2/5}$	$p^{1/10}q^{1/2}$
6	$q^{1/3}$	$q^{1/3}$	$p^{1/15}q^{2/5}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
l	$q^{2/l}$	$q^{2/l}$	$p^{2/(l^2-l)}q^{2/(l-1)}$

TABLE I: In column one, the crossover temperature $T_\times$ from an "ordered" to a "disordered" state is determined by defining the ratio $\omega_l = 1$ of the sum of Boltzmann weights of sub-optimal node assignments to the weight of the optimal assignment into clique communities as a function of the cluster size l and the number of communities q in the large q limit. In column two, we estimate $T_\times^{(1/q)} \simeq T_\times$ through different means by calculating the probability $p = 1/q$ that two nodes (in the same clique ideally) are determined to be in the same cluster. In the last column, we generalize column two for an arbitrary probability p.

where σ_i and σ_j denote the cluster memberships for nodes i and j, respectively. Expressing the numerator and in terms of l and q, Eq. (13) becomes,

$$\begin{aligned} P(\sigma_i = \sigma_j) &= \Big\{ q e^{2\beta\binom{l}{2}} + (l-2)q(q-1)e^{2\beta\binom{l-1}{2}} \\ &+ \cdots + q(q-1)\cdots(q-l-2) \Big\} \\ &\Big/ \Big\{ q e^{2\beta\binom{l}{2}} + lq(q-1)e^{2\beta\binom{l-1}{2}} \\ &+ \cdots + q(q-1)(q-2)\cdots(q-l) \Big\} \end{aligned} \tag{14}$$

In the limit of large q, Eq. (14) simplifies to

$$P(\sigma_i = \sigma_j) \simeq \frac{q e^{2\beta\binom{l}{2}} + \sum_{k=1}^{l-2} \binom{l-2}{k} q^{k+1} e^{2\beta\binom{l-k}{2}}}{q e^{2\beta\binom{l}{2}} + \sum_{k=1}^{l} \binom{l}{k} q^{k+1} e^{2\beta\binom{l-k}{2}}}. \tag{15}$$

Choosing $P(\sigma_i, \sigma_j) = 1/q$ yields in a crossover temperature $T_\times^{(1/q)}$ at which the system goes from being unbroken q-state symmetry to ordered. When $l = 3$, Eq. (15) becomes,

$$q^2 e^{6\beta} + q^3 e^{2\beta} \simeq q e^{6\beta} + 3q^2 e^{2\beta} + 3q^3 + q^4. \tag{16}$$

In the large q limit, $e^{2\beta} \simeq q^{2/3}$, and the crossover temperature is $T_\times^{(1/q)} = 3/\log q$. The asymptotic expressions for several values of q and l are summarized in column two of Table I. For general q and l, the relation is

$$T_\times^{(1/q)} \simeq \frac{l}{\log q}. \tag{17}$$

Eq. (17) is consistent with Eq. (12) and also [59].

For a general crossover probability $P(\sigma_i, \sigma_j) = p$ with $l = 3$, the crossover temperature $T_\times^{(p)}$ is determined by solving

$$e^{6\beta} + q e^{2\beta} \simeq p e^{6\beta} + 3qpe^{2\beta} + 3q^2 p + q^3 p. \tag{18}$$

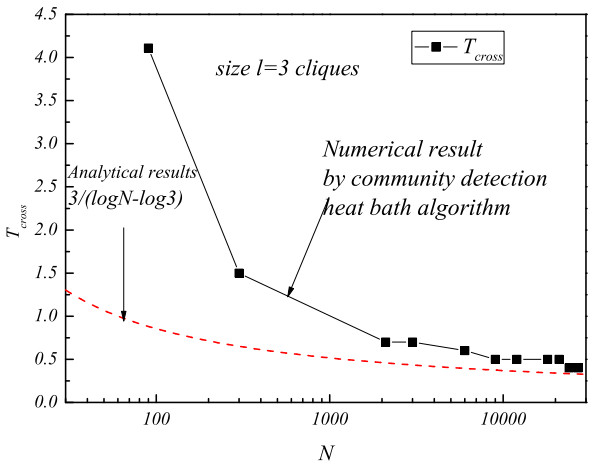


FIG. 18: (Color online) The crossover temperature at which the system cannot be perfectly solved as the function of the system size N. The data here uses cliques of size $l = 3$. The dashed line is the analytical result and the solid line is determined by the heat bath community detection algorithm optimizing the Hamiltonian of Eq. (1).

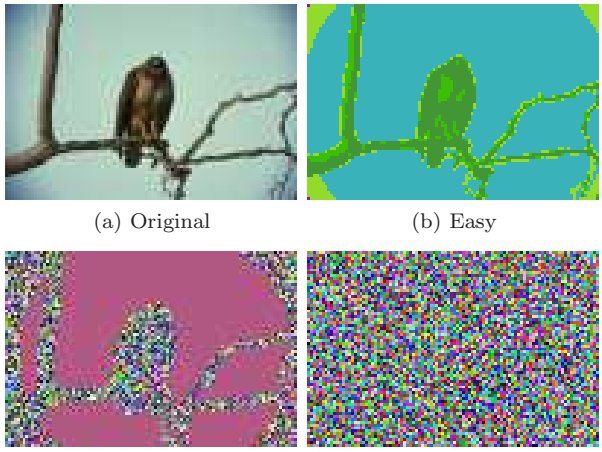


FIG. 19: (Reproduced from Ref. [60]) We show an image where we apply our community detection algorithm to detect the relevant structures. This case seeks to identify a bird and tree against a sky background. The original images is in panel (a), and the segmentation results are shown in panels (b–d) corresponding to the easy, hard, and unsolvable regions of the community detection problem, respectively. Figure 20 shows the phase diagram identifying these respective regions.

In the large q limit, Eq. (18) is $e^{2\beta} \simeq p^{1/3}q$, where $T_\times^{(p)} \simeq 2/(\log q + 1/3 \log p)$. Results for $T_\times^{(p)}$ for several values of q and l are shown column three of Table I. For general q and l, the relation is

$$T_\times^{(p)} \simeq \frac{1}{p^{1/l}q}. \tag{19}$$

C. Simulated crossover temperature

We can also simulate the crossover temperature $T_\times$ or $T_\times^{(p)}$ as a function of system size N by solving the non-interacting clique problem using our heat bath community detection algorithm (see Appendix D). As seen in Fig. 18, the simulated and analytic asymptotic behaviors agree well in the large N limit, so the crossover temperature for this trivial system is $T_\times = 0$.

The crossover temperature derived in this section deals with a *heat-induced* disorder. That is, it marks the onset of a "liquid" phase that transitions at a lower heat bath temperature as the system size grows. In practice, one uses a SA algorithm that applies a cooling scheme (as opposed to a constant temperature HBA) to improve the attempt at locating the ground state of the system. That is, it applies a high temperature exploration of the general landscape finished by low temperature "fine tuning" of the solution. For the non-interacting cliques in this section, SA would obviously still identify the ground state because the energetic fluctuations would trivially diminish as the system is cooled toward $T = 0$.

With increasing p_{out} at low T, disorder imposed by the glass-type transition is induced by the complexity of the energy landscape, but the transition is qualitatively comparable in the sense of the induced disorder in the solutions found by the HBA. The glass phase also experiences a transition to a liquid-like disordered state at a temperate that increases slowly with the level of noise, but here, a SA solver will not necessarily transition readily to the ideal solution as the system is cooled because of the inherent complexity of the energy landscape. The greedy algorithm used in [34] (equivalent to the HBA at $T = 0$) applied to the Potts model of Eq. (1) is already very accurate [11, 18, 34], so we expect that the greatest benefit of SA over a greedy-oriented solver using Eq. (1) will manifest in the hard region near the onset of the "glassy" transition.

D. A discussion of the crossover temperature

For a spin system with fixed size N, a larger number of spin states q corresponds to a more disordered system. If we expand the partition function of the Potts model in terms of $1/q$, it is explicitly represented as a sum over configurations with progressively larger clusters of identical spins [53]. That is, two spins with the same index $\sigma_i = \sigma_j$ are connected. Then three spins $\sigma_i = \sigma_j = \sigma_k$ are connected, etc. The resulting terms illustrate that increasing q emulates increasing temperature T.

Our analysis in this section applies to general graphs with ferromagnetic interactions (equivalent to the "label propagation" community detection algorithm [54]) on regular, fixed-coordinate lattices [55–57]. Increasing the number of system states q causes the system to be in-

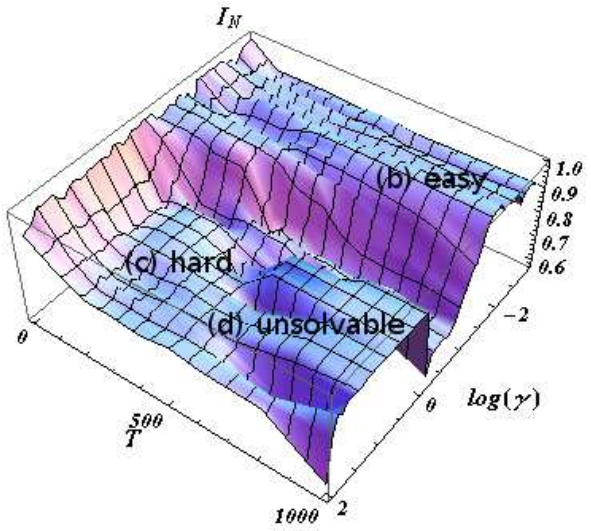


FIG. 20: (Reproduced from Ref. [60]) We show a three-dimensional phase diagram of NMI (I_N) versus $\log(\gamma)$ and $\log(T)$ for the image segmentation of the bird in Fig. 19. T is heat bath temperature for a stochastic community detection solver (see Appendix D), and γ is the model weight in Eq. (1). We note that the optimal values in the easy and hard regions correspond to the "physical" segmentations of the bird and tree against the background, but the bird is undetectable in the unsolvable region.

creasingly disordered. Thus, in the community detection problem, increasing number of communities q linearly with the system size N (such that the average community size remains constant), the solvable (easy) phase shrinks to a "small" region as $N \to \infty$.

Figures 13(m–p) illustrate the distinction in the different regions or types of disorder: entropic (high complexity) and energetic (high T). Interestingly, in some cases, additional noise emulates a higher temperature solution process in the sense that it provides additional avenues to explore different configurations. Such an effect may occur in Fig. 13(a-d) where the accuracy [I_N in panel (b) *increases* for a short time with increasing noise p_{out}].

Fig. 13(n) further shows a crossover region $0.24 \lesssim p_{out} \lesssim 0.32$ where mid-range temperatures improve the solution accuracy (higher I_N). Although this data uses a constant temperature heat bath (no cooling schedule), this is the effect of a stochastic solver (see Appendix D), allowing it to navigate the difficult energy landscape more accurately than a greedy solver. On the left (lower T), the more greedy nature of the solver prevents an accurate solution in the presence of high noise. On the right, the higher temperature of the heat bath itself hinders an accurate solution. In effect, the HBA "wanders" at energies above the meaningful, but locally complex, features of the energy landscape resulting in more random solutions.

The results here incorporate a "global" model parameter γ in Eq. (1). That is, the model asserts globally optimal γ(s) for the entire graph. For large graphs, this condition is less likely to be true across the full scope of the network, but one can explore methods to obtain locally optimal γ_ℓ (in time or space) for each region or cluster ℓ [58]. Utilizing locally optimal γ_ℓs will likely work to circumvent the temperature transition at low levels of noise. The successful selection of a local γ_ℓ in the glassy (high noise) region is more difficult because of the complex nature of the local energy landscape.

In the following section, we study the free energy of several systems for ferromagnetic Potts models and then generalize to *arbitrary* weighted Potts models, including antiferromagnetic interactions, on *arbitrary* graphs [59].

VI. AN EXAMPLE OF A PHASE TRANSITION IN AN IMAGE SEGMENTATION PROBLEM

We illustrate the phase transition effect with an realistic image segmentation example [60]. In Fig. 19, we apply our community detection algorithm to detect a bird and tree against a sky background. We display the results in Fig. 20 where we plot NMI (I_N) versus $\log(\gamma)$ in Eq. (1) and $\log(T)$ where T is the temperature for our stochastic community detection solver (see Appendix D). For this problem, we apply edge weights by replacing the A_{ij} elements in Eq. (1) with "attractive" and "repulsive" weights w_{ij} which are defined by regional intensity differences within the image [60].

We label the easy (b), hard (c), and unsolvable (d) regions in the phase plot for the bird image in panel (a). Panel (b) shows that our algorithm clearly detects the bird and tree against the background, meaning that the NMI information measure identifies the physically relevant clusters in the problem. In panel (c), the background is segmented separately, but the bird and tree are composed of many small clusters. Panel (d) shows that the bird is undetectable in the unsolvable region.

VII. FREE ENERGY: SIMPLE RESULTS

In the following analysis, we explicitly show the large q and large T expansions for the free energy per site in three example systems (a non-interacting clique system, simple interacting clique system, and a random graph) before generalizing the analysis to arbitrary unweighted and weighted graphs. Previous works examined disorder transitions for random-bond Potts models [61, 62] and Ref. [63] studied zeros of the partition function in the large q limit. Large q behavior was shown to approach mean-field theoretical results on fixed lattices [64, 65]. For the unweighted systems, we use a binary distribution for the interaction strength $J = 1$ or 0 (*i.e.*, the energy contribution of an edge is either "on" or "off").

A. Free energy of a non-interacting clique system under a large q expansion

If we generalize the non-interacting clique system in Fig. 17 to cliques of size l, the partition function is

$$\begin{aligned}\mathbf{Z} \;=\; \Bigg[& q e^{\beta J \binom{l}{2}} + lq(q-1) e^{\beta J \binom{l-1}{2}} \\ & + \frac{l(l-1)}{2} q(q-1)(q-2) e^{\beta J \binom{l-2}{2}} \\ & + \cdots + q(q-1)(q-2)\cdots(q-l) \Bigg]^q . \end{aligned} \tag{20}$$

When $q \to \infty$,

$$\begin{aligned}\mathbf{Z} \;\approx\; \Bigg[& q e^{\beta J \binom{l}{2}} + lq^2 e^{\beta J \binom{l-1}{2}} + \frac{l(l-1)}{2} q^3 e^{\beta J \binom{l-2}{2}} \\ & + \cdots + q^{l+1} \Bigg]^q . \end{aligned} \tag{21}$$

The free energy per site, $f = -\frac{k_B T}{N} \log \mathbf{Z}$, (with the Boltzmann constant set to $k_B = 1$) is

$$f \approx -T \log q - T \sum_{k=0}^{l-2} \left[\binom{l-1}{k} \frac{1}{k+1} e^{\beta J \binom{l-1}{k-1}} \right] q^{-(k+1)} . \tag{22}$$

From Eq. (22), we further simply the free energy per site

$$\begin{aligned} f \;&\approx\; -T \log q - T \sum_{k=0}^{l-2} a(k) e^{\beta J \binom{l-1}{k-1}} q^{-(k+1)} \\ f \;&\approx\; -T \log q - T a(0) \frac{e^{\beta J}}{q} \end{aligned} \tag{23}$$

where $a(k) = \binom{l-1}{k} \frac{1}{k+1}$. We will compare Eq. (23) with the high T expansion in the next section. Despite the functional dependence of $\exp(\beta J)$, the large q limit dominates the expansion, forcing the system to be approximately equivalent to a large temperature limit.

B. Free energy of a non-interacting clique system as ascertained from a high temperature expansion

Note that the *most* ordered Potts graph is a system of non-interacting cliques (maximally connected subgraphs). That is, the presence of noise (extraneous inter-community edges) will only serve to *increase* the overall disorder in the system. One exception is that increased disorder can emulate increased temperature T for both greedy and stochastic community detection solvers (see also Sec. V D).

We can construct the high T expansion easily by means of Tutte polynomials [66] (see Appendix E 1) where we again solve a system of q cliques of size l. Equation (1) and a ferromagnetic Potts model have the same *ground state* energy for this clique system (see also Secs. VII E, VII G, and VII H for more general derivations), so the partition function in terms of the Tutte polynomial $t(G; x, y)$ for a graph G is

$$\mathbf{Z} = q^{k(G)} v^{|V| - k(G)} t(G; x, y) \tag{24}$$

where q is the number of clusters or states, $v = \exp(\beta J) - 1$, G denotes the graph, $k(G)$ is the number of connected components in G, $|V| = N$ is the number of vertices, $x = (q+v)/v$ and $y = v + 1$. For the non-interacting clique system, $k(G) = q$ and $N = lq$. We denote the Tutte polynomial of a single clique of size l as $K_l(G; x, y)$.

$K_2(G; x, y) = x$, so the partition function is

$$\begin{aligned} \mathbf{Z} \;&=\; q^q v^q x^q , \\ \mathbf{Z} \;&=\; q^q v^q \left(\frac{q}{v} + 1 \right)^q , \end{aligned} \tag{25}$$

where we used $N = 2q$. In a high T approximation, $x \approx q/v \gg 1$, so the partition function becomes $Z \approx q^{2q}$, and the free energy is

$$f \approx -T \log q, \tag{26}$$

which simply states that the system is completely random in the large T limit.

For triangle cliques, $K_3(G; x, y) = x^2 + x + y$. The graph G is composed of disjoint triangles, so the Tutte polynomial is $t(G; x, y) = \left(x^2 + x + y \right)^q$, and the partition function becomes

$$\mathbf{Z} \approx q^q v^{2q} \left(x^2 + x + y \right)^q . \tag{27}$$

In a high T approximation $y \approx 1$, but $x \approx q/v \gg 1$ in either the large q or large T limits, so we make a further approximation of $y \approx 0$. Then, $K_3(G; x, y = 0) = x^q (x+1)^q \approx x^{2q}$. The partition function simplifies to $\mathbf{Z} \approx q^{3q}$, so the free energy per site for $l = 3$ is again

$$f \approx -T \log q \tag{28}$$

which is identical to the $l = 2$ result because we consistently applied the approximation $q/v \gg 1$ to $x = (q/v + 1) \approx q/v$ and $(x + 1) = (q/v + 2) \approx q/v$.

Generalizing to an arbitrary clique size l in the high T approximation, the Tutte polynomial $K_l(G; x, y = 0)$ is

$$K_l(G; x, y = 0) = \frac{\Gamma(x + l - 1)}{\Gamma(x)}, \tag{29}$$

The partition function is

$$\mathbf{Z} \approx q^{lq} \left(\frac{v}{q} \right)^{(l-1)q} \frac{\Gamma\left(\frac{q}{v} + l - 1 \right)}{\Gamma\left(\frac{q}{v} \right)}, \tag{30}$$

and $v = e^{\beta J} - 1 \approx \beta J$, so the free energy per site yields

$$f \approx -T \log q - \frac{l-1}{l} T \log \left(\frac{\beta J}{q} \right) - \frac{T}{lq} \log \left[\frac{\Gamma\left(\frac{q}{\beta J} + l - 1 \right)}{\Gamma\left(\frac{q}{\beta J} \right)} \right] . \tag{31}$$

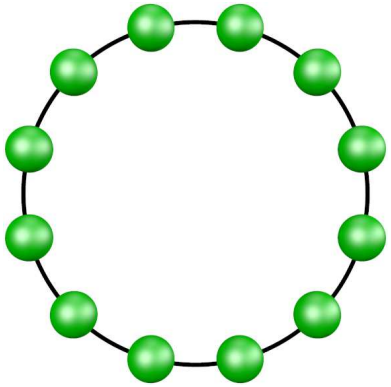


FIG. 21: (Color online) A depiction of a circle of cliques (maximally connected clusters) of size l connected by single edges. In contrast to Fig. 17, this system adds a simple interaction between cliques. We analyze the configuration in Sec. VII C and show that a ferromagnetic Potts model behaves the same in the large q and large T limits.

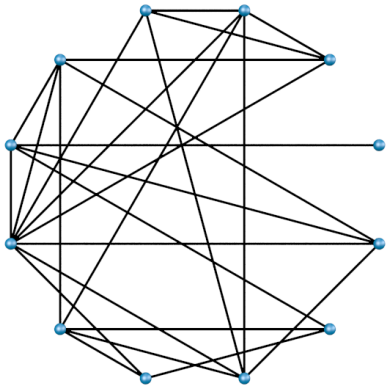


FIG. 22: (Color online) A sample depiction of a random graph with N nodes. In Sec. VII D, we analyze such a system by randomly removing edges from a clique configuration of N nodes under the assumption that we maintain a connected graph. We show that a ferromagnetic Potts model on a random graph behaves the same in the large q and large T limits.

The leading $\log q$ term represents the infinite T limit which is approximately constant in large systems for any clique size l. That is, the partition function $\mathbf{Z}_{T\to\infty} \approx q^N$ for every system. The $l = 2$ and 3 results above illustrate that when $l \ll q$, the ratio of gamma functions in Eq. (31) simplifies to x^{lq}, and the free energy for the non-interacting clique system is approximately $f \approx \log q$ in the large T limit.

The second term in Eq. (31) gives the leading order correction for high T. It is absent in the explicit $l = 2$ and 3 results above because we applied the approximation $q/v \gg 1$. Together, the last two terms imply that increasing the temperature T (decreasing β) emulates increasing the number of communities q for a ferromagnetic Potts model.

C. Free energy for the "circle of cliques" in the high q or the high T expansion

We now investigate the slightly more complicated system depicted in Fig. 21: a "circle of cliques" where each complete sub-graph cluster is connected to its neighbors by a single edge. We construct q cliques of size $l = 3$ and apply the Tutte polynomial method [66] to solve the system. As in the previous sub-section, the ground state of Eq. (1) and a ferromagnetic Potts model have the same energy, so we use a ferromagnetic model. In terms of the Tutte polynomial $t(G; x, y)$ for a graph G, the partition function is given by Eq. (24).

Equation (F4) in Appendix F derives the exact Tutte polynomial for Fig. 21 with $l = 3$, and Eq. (F6) gives the high T expansion $t(G; x, y = 0) = (1 + x)^{q+1} x^{2q-3}$. Substituting $N = 3q$ and the approximation $x \approx q/v$ (in either the large q or large T limits), the partition function becomes

$$\mathbf{Z} \approx q^{2q-2} v^{q+2} (1 + x)^{q+1} x^{2q-3} \tag{32}$$

We factor out q^{3q}, and then apply the approximations: $v = \exp(\beta J) - 1 \approx \beta J$, $x \approx q/v \approx q/(\beta J) \gg 1$, and $q \gg 1$. The free energy per site is then

$$f \approx -T \log q - \frac{2T}{3} \log\left(\frac{q}{\beta J}\right) \tag{33}$$

As in the previous sub-section, the leading $\log q$ term represents the infinite T limit. Equation (33) affirms the implication of Eq. (31) regarding the corresponding behavior of large q or T. Specifically, increasing the temperature (decreasing β) emulates increasing the number of communities q for a ferromagnetic Potts model.

D. Free energy of a random graph in a large q or a large T expansion

We apply the Tutte polynomial method of Appendix E 1 to determine the high T and high q partition function for a random graph. For calculation purposes, we begin with a complete graph of size N. Then we randomly remove edges to construct a random graph such that any two nodes are connected by and edge with a probability p. The derivation repeatedly applies lemma 1 stated in Appendix E 1.

We denote the Tutte polynomial of a complete graph (clique) of size l as K_l. $t(G)$ for a clique with d duplicated edges (multiply defined edges between two nodes) or loops (self-edges) is defined as $K_l^{(d)}$. For economy of notation, we also define $G_l^{[m]}$ as the Tutte polynomial of a graph with m *missing edges* (*i.e.*, not a clique). Note that $K_l^{(0)} \equiv G_l^{[0]} \equiv K_l$. For the following derivation,

we work under the *assumption* that when we delete or contract any edge, the random graph remains connected.

Under the high temperature T or high number of clusters q approximations, $y \ll x$ and $y \simeq 0$. Equation (29) gives the exact expression of the Tutte polynomial $K_l(G; x, y = 0)$ for a clique at $y = 0$. If we cut one edge from the complete graph K_N, we obtain the recursion formula

$$\begin{aligned} K_N &= G_N^{[1]} + K_{N-1}^{(N-1)}, \\ K_N &= G_N^{[1]} + K_{N-1}. \end{aligned} \tag{34}$$

where we applied lemma 1 to obtain Eq. (34). From henceforth, we assume the application of lemma 1. We are interested in the graph with missing edges, so we solve Eq. (34) for $G_N^{[1]}$.

$$G_N^{[1]} = K_N - K_{N-1}. \tag{35}$$

Note that the reduced graph is represented as a summation over complete graphs.

Now we apply the Tutte recursion formula to *both sides* of Eq. (35).

$$G_N^{[2]} + G_{N-1}^{[1]} = G_N^{[1]} + K_{N-1} - G_N^{[1]} - K_{N-1}. \tag{36}$$

We can choose the deleted and contracted edges in the corresponding terms to be identical because the resulting Tutte polynomial is in general independent of the operation order. After collecting terms and substituting the previous $G_N^{[1]}$ result, we solve for $G_N^{[2]}$ to obtain

$$G_N^{[2]} = K_N - 2K_{N-1} + K_{N-2}, \tag{37}$$

for this particular random graph. Again, the right-hand-side of Eq. (37) is a summation over complete graphs. This recursive relation for $G_N^{[k]}$ continues until we obtain

$$G_N^{[k]} = \sum_{i=0}^{k} (-1)^i \binom{k}{i} K_{N-i}. \tag{38}$$

We insert this into Eq. (29) with the pre-factor qv^{N-1} to generate the partition function at high T

$$\mathbf{Z} = q^N \left(\frac{v}{q}\right)^{N-1} \left[\sum_{i=0}^{k} (-1)^i \binom{N}{i} \frac{\Gamma(x+N-i-1)}{\Gamma(x)}\right]. \tag{39}$$

We substitute $x = (q+v)/v \approx q/v$ when $v \ll q$ (high T or high q approximations) and again utilize $v = e^{\beta J} - 1 \approx \beta J$ in the high T approximation to obtain the free energy per site

$$\begin{aligned} f &= -T \log q - \frac{N-1}{N} T \log\left(\frac{\beta J}{q}\right) \\ &\quad - \frac{T}{N} \log\left[\sum_{i=0}^{k} (-1)^i \binom{N}{i} \frac{\Gamma\left(\frac{q}{\beta J} + N - i - 1\right)}{\Gamma\left(\frac{q}{\beta J}\right)}\right] \end{aligned} \tag{40}$$

Note that the first two terms become $\log(q)/N \log(\beta J)$ as $N \to \infty$. From Eq. (40), we obtain the same conclusion for this random graph as for the previously analyzed clique systems. While Secs. VII A, VII B, and VII C result in free energies with different functional forms, in each case, q and T have the same functional form in the arguments of the functions in the high T limit.

E. Free energy of an arbitrary graph G in the large T expansion

We can construct the explicit high T expansion for an arbitrary (unweighted) graph G by means of the Tutte polynomial method [66]. Factoring out q^N and substituting $|V| = N$, $x = q/v + 1$, and $y = v + 1$ in Eq. (24), we write a trivially modified form of the partition function

$$\mathbf{Z} = q^N \left[\left(\frac{v}{q}\right)^{N-k(G)} t\left(G; \frac{q}{v} + 1, v + 1\right)\right]. \tag{41}$$

At this point, the equation is completely general, but the corresponding behavior for temperature T and number of clusters q is almost apparent in the reciprocal relationship of q and v.

Again, $x \approx q/v$ in either the large q or large T limits. In a high T approximation, $v \approx \beta J = T/J$ and $y \approx 0$ or 1 ($y = 0$ is a common approximation since $x \gg y$ in the same limit).

$$\mathbf{Z} \approx q^N \left[\left(\frac{J}{qT}\right)^{N-k(G)} t\left(G; \frac{qT}{J}, y_{_{T'}}\right)\right] \tag{42}$$

where $y_{_{T'}} = 0$ or 1. The free energy per site is then

$$f \approx -T \log q - \frac{N-k(G)}{N} T \log\left(\frac{J}{qT}\right) - \frac{T}{N} \log\left[t\left(\frac{qT}{J}, y_{_{T'}}\right)\right]. \tag{43}$$

The leading $\log q$ term appears in our previous calculations. Again, it represents the infinite T limit for an arbitrary system which is approximately constant in large systems.

From the perspective of increasing q, the similarity to the large T behavior is more apparent if we fix the temperature $T = T'$ and define an effective interaction constant $J_q \equiv e^{J/T'} - 1$. We then rewrite Eq. (43) as

$$f \approx -T \log q - \frac{N-k(G)}{N} T \log\left(\frac{J_q}{q}\right) - \frac{T}{N} \log\left[t\left(\frac{q}{J_q}, y_q\right)\right]. \tag{44}$$

where $y_q \equiv e^{J/T'}$ is a constant. When $N \to \infty$ and $k(G) \ll N$, the first two terms become $T \log(q)/N \log(\beta J)$. Comparing Eqs. (43) and (44) shows the close correspondence between increasing q (at fixed T') and increasing T. J_q grows exponentially faster than q with decreasing T', so a finite (perhaps small) stable or solvable region is likely except in the presence of high noise.

F. Annealed versus quenched averages

The above proofs apply to quenched averages because the binary distribution is constant with respect to the distribution integration. That is, using Eq. (44), we assume a probability distribution $P(\{J_{ij}\})$ and integrate over it to obtain the quenched average free energy per site

$$\begin{aligned} f[\{J_{ij}\}] &= \int DJ_{ij} \prod_{i\neq j} P(\{J_{ij}\}) \Bigg\{ \log q \\ &+ \frac{N-k(G)}{N} \log\left(\frac{J}{qT}\right) \\ &+ \frac{1}{N} \log\left[t\left(\frac{qT}{J}, y_{T'}\right)\right] \Bigg\}, \end{aligned} \tag{45}$$

but the integrand (f_0) is a constant because J is independent of $\{J_{ij}\}$, so the integral trivially simplifies to

$$f[\{J_{ij}\}] = f_0 \int DJ_{ij} \prod_{i\neq j} P(\{J_{ij}\}). \tag{46}$$

where the integral is unity. In a more general model with a defined $\{J_{ij}\}$ probability distribution, the leading order $\log q$ contribution would remain unchanged, but we would obtain correction terms from the integration over the quenched interaction distribution $\{J_{ij}\}$.

G. Free energy of non-interacting cliques for an *arbitrary* weighted Potts model under a large T expansion

We can represent an *arbitrary* weighted Potts model with ferromagnetic and antiferromagnetic interactions. That is, we can generally write

$$H(\{\sigma\}) = -\frac{1}{2} \sum_{i\neq j} \left[a_{ij} A_{ij} - b_{ij}(1-A_{ij})\right] \delta(\sigma_i, \sigma_j). \tag{47}$$

where a_{ij} and b_{ij} are arbitrary "attractive" and "repulsive" edge weights. This summarization includes modularity [50], a Potts model incorporating a "configuration null model" (CMPM) comparison [9] (the most common variation in [9] is effectively generalizes modularity), CMPM allowing antiferromagnetic relations [67], "label propagation" [14, 54], an Erdős-Rényi Potts model [9, 68], a "constant Potts model" [49], the weighted form of the APM [11, 34], or a "variable topology Potts model" suggested in [11].

Note that the repulsive weights b_{ij} are important in that they provide a "penalty function" which enables a well-defined ground state for the Hamiltonian for an arbitrary graph. That is, the ground state of a purely ferromagnetic Potts model in an arbitrary graph is trivially a fully collapsed system (perhaps with disjoint subgraphs). Several of the above models incorporate a weighting factor γ of some type on the penalty term which allows the model to span different scales of the network in qualitatively similar ways.

We denote a the partition function of a graph G^* with l nodes and weighted edges $\{e\}$ by $Z(G^*; q, \mathbf{v}) \equiv \mathcal{K}_l$. We assume that $J_e \ll T$ for all edges e, and all pairs of nodes in G^* are connected by a weighted edge J_e (either ferromagnetic or antiferromagnetic). From Appendix E 2, a recurrence relation for the multivariate Tutte polynomial of a general weighted clique is

$$\mathcal{K}_l \approx \left(q + \sum_{k=1}^{j-1} v_{k+\ell_j}\right) \mathcal{K}_{l-1} + O(y_e), \tag{48}$$

The partition function for $\mathcal{K}_l$ at high T is

$$\mathcal{K}_l \approx q^N \prod_{j=2}^{l} \left(1 + \sum_{k=1}^{j-1} \frac{v_{k+\ell_j}}{q}\right), \tag{49}$$

Now, we generate a graph consisting of a set of q non-interacting cliques of size l_i where $i = 1, 2, \ldots, q$.

$$\mathcal{K}_l \approx q^N \prod_{i=1}^{q} \prod_{j=2}^{l_i} \left(1 + \sum_{k=1}^{j-1} \frac{\beta J_{k+\ell_j}}{q}\right). \tag{50}$$

where we used $v_e \approx \beta J_e$ at high T for general edge weights J_e (even if $J_e < 0$ as long as $J_e \ll T$).

The free energy is

$$\begin{aligned} f &\approx -T\log q - \frac{T}{N} \sum_{i=1}^{q} \sum_{j=2}^{l_i} \sum_{k=1}^{j-1} \frac{\beta J_k}{q} \\ &\approx -T\log q - \frac{1}{N} \sum_{i=1}^{q} \frac{E_i}{q} \\ &= -T\log q - \frac{E}{qN} \end{aligned} \tag{51}$$

where we invoked $\log(1+x) \approx x$ for $x \ll 1$ there. E_i is the energy of cluster i according to the weighted Potts model of Eq. (47), and E is the total energy of the graph. Equations (50) and (51) both imply that large q emulates large T for an *arbitrary* Potts model on a weighted graph G. That is, if a community detection quality function can be expressed in terms of the general Potts model in Eq. (47), then large q and large T are essentially equivalent.

H. Free energy of non-interacting cliques for an *arbitrary* weighted Potts model under a large q expansion

The multivariate Tutte polynomial [69] (see also Appendix E 2 and Ref. [59]) appears in a subgraph expansion over the subset of edges $\mathcal{A} \subseteq \mathcal{E}$ in a graph $G = (V, \mathcal{E})$ with a set of V vertices and $\mathcal{E}$ edges

$$Z(G; q, \mathbf{v}) = q^N \left(1 + \sum_{e'=1}^{|\mathcal{E}|} \frac{v_{e'}}{q} + \cdots + q^{k(G)-N} \prod_{f'=1}^{|\mathcal{E}|} v_{f'}\right) \tag{52}$$

$k(\mathcal{A})$ is the number of connected components of $G_A = (V, \mathcal{A})$ and $v_e = \exp(\beta J_e) - 1$. For our purposes, Eq. (52) serves as an alternate representation of Z_G to facilitate the calculation of the large q expansion.

For large q, when $q^N \gg |v_e|^L$, we may neglect the last term, and for a system of non-interacting cliques of sizes l_i with $i = 1, 2, \ldots, q$, the leading order terms in large q are

$$Z(G; q, \mathbf{v}) \approx q^N \prod_{i=1}^{q} \prod_{j=1}^{n_i} \left(1 + \sum_{k=1}^{j-1} \frac{v_{k+\ell_j}}{q} \right) \tag{53}$$

The approximation is identical to Eq. (49) at high T. Ref. [59] calculates an explicit crossover temperature including the last subgraph $\mathcal{A} = \mathcal{E}$ that competes with the large q terms as $T \to 0$. The free energy corresponding to Eq. (53) becomes

$$f \;\approx\; -T \log q - \frac{T}{N} \sum_{i=1}^{q} \sum_{j=2}^{l_i} \sum_{k=1}^{j-1} \frac{v_k}{q} \tag{54}$$

where we applied the small x approximation $\log(1+x) \approx x$.

In order to illustrate the correspondence in large q and T, we fix $T = T'$, define $J_e^{(q)} \equiv \exp(\beta' J_e) - 1$, and rewrite the free energy per site

$$f \;\approx\; -T' \log q - \frac{T'}{N} \sum_{i=1}^{q} \sum_{j=2}^{l_i} \sum_{k=1}^{j-1} \frac{J_k^{(q)}}{q}. \tag{55}$$

Large q in Eq. (52) emulates large T in Eq. (50). As with the unweighted case in Eq. (44) in Sec. VII E, $J_e^{(q)}$ is exponentially weighted in $\beta' = 1/T'$, so a non-zero (perhaps small) region of stability is essentially ensured except in the presence of high noise [59]. We can additionally determine a rigorous bound using methods in [59, 70, 71]

$$T_\times^{\mathrm{UB}} = \frac{\bar{J}_0}{\log\left[\frac{p(q-1)}{(1-p)}\right]}, \tag{56}$$

where $\bar{J}_0 = \frac{1}{2} \sum_j J_{j0} \left[1 + \mathrm{sgn}(J_{j0})\right]$ is a generous upper bound summing only positive energy contributions and p is the probability for finding a given spin σ_0 in a specific spin state $\bar{\sigma}$. This result further agrees with our conclusions. Note that as $p \to 1/q$, the system is completely disordered, so $T_\times \to \infty$. As $p \to 1$, the system is perfectly ordered, so $T \to 0$.

VIII. CONCLUSIONS

We systematically examined the phase transitions for the community detection problem via a "noise test" across a range of parameters. The noise test consists of a structured graph with a strongly-defined ground state. We add increasing numbers of extraneous intercommunity edges (noise) and test the performance of a stochastic community detection algorithm in solving for the well-defined ground state. Specifically, we studied two types (sequences) of systems. In the first such sequence of systems in Fig. 5, we fixed the ratio $\alpha = q/N$ of the number of communities q to the number of nodes N. We fixed q at different values and varied N in the second sequence of systems in Fig. 9. In Fig. 13, we explored the largest tested systems with $N = 2048$ nodes in more detail where we depicted additional measures to illustrate the transitions. All of these systems showed regions with distinct phase transitions in the large N limit. Deviations occurred most often in smaller systems indicating a definite finite-size effect.

The spin-glass-type phase transitions in our noise test occurred between solvable and unsolvable regions of the community detection problem. A hard, but solvable, region lies at the transition itself where it is difficult, in general, for any community detection algorithm to obtain the correct solution. We analyzed a system of non-interacting cliques and illustrated that in the large q limit, the system experiences a thermal disorder in the thermodynamic limit for any non-zero temperature. When in contact with a heat bath, the asymptotic behavior of the temperatures beyond which the system is permanently disordered varies slowly with the number of communities q, specifically, $T_\times \simeq O[1/\log q]$. This implies that problems of practical size maintain a definite region of solvability. Given the connection between Jones polynomials of knot theory and Tutte polynomials for the Potts model, our results imply similar transitions in large random knots (see Appendix G).

We further studied the free energy of arbitrary graphs arriving at the same conclusion. Increasing number of communities q emulates increasing T in arbitrary graphs for a general Potts model. The effective interaction strength for increasing q scales such that this disorder is circumvented by the often standard use of a simulated annealing algorithm, but the "glassy" (high noise) region remains a challenge for any community detection algorithm.

Acknowledgments

This work was supported by NSF grant DMR-1106293 (ZN). We also wish to thank S. Chakrabarty, R. Darst, P. Johnson, V. Dobrosavljevic, B. Leonard, A. Middleton, M. E. J. Newman, D. Reichman, V. Tran, and L. Zdeborová for discussions and ongoing work.

Appendix A: Definitions: Trials and Replicas

We review the notion of trials and replicas on which our algorithms are based. Both pertain to the use of multiple identical copies of the same system which differ from

one another by a permutation of the site indices. Thus, whenever the time evolution may depend on sequentially ordered searches for energy lowering moves (as it will in our greedy algorithm), these copies may generally reach different final candidate solutions. By the use of an *ensemble* of such identical copies (see, e.g., Fig. 2), we can attain accurate result as well as determine information theory correlations between candidate solutions and infer from these a detailed picture of the system.

In the definitions of "trials" and "replicas" given below, we build on the existence of a given algorithm (any algorithm) that may minimize a given energy or cost function. In our particular case, we minimize the Hamiltonian of Eq.(1).

• *Trials.* We use trials alone in our bare community detection algorithm. We run the algorithm on the same problem t independent times. This may generally lead to different contending states that minimize Eq.(1). Out of these t trials, we will pick the lowest energy state and use that state as the solution.

• *Replicas.* We use both trials and replicas in our multiscale community detection algorithm. Each sequence of the above described t trials is termed a *replica*. When using "replicas" in the current context, we run the aforementioned t trials (and pick the lowest solution) r independent times. By examining information theory correlations between the r *replicas* we infer which features of the contending solutions are well agreed on (and thus are likely to be correct) and on which features there is a large variance between the disparate contending solutions that may generally mark important physical boundaries. We will compute the information theory correlations within the ensemble of r replicas. Specifically, *information theory extrema* as a function of the scale parameters, generally correspond to more pertinent solutions that are locally stable to a continuous change of scale. It is in this way that we will detect the important physical scales in the system (see Fig. 2).

Appendix B: Information theory and complexity measures

We use information theory measures to calculate correlations between community detection solutions and expected partitions in the noise test problem. To begin, N nodes of partition A are partitioned into q_A communities of size $\{n_a\}$ where $1 \leq a \leq q_A$. The ratio n_a/N is the probability that a randomly selected node is found in community a. The Shannon entropy is

$$H_A = -\sum_{a=1}^{q_A} \frac{n_a}{N} \log_2 \frac{n_a}{N} \tag{B1}$$

The mutual information $I(A,B)$ between partitions A and B is

$$I(A,B) = \sum_{a=1}^{q_A} \sum_{b=1}^{q_B} \frac{n_{ab}}{N} \log_2 \frac{n_{ab} N}{n_a n_b} \tag{B2}$$

where n_{ab} is the number of nodes of community a in partition A that are also found in community b of partition B. The normalized mutual information $I_N(A,B)$ is then

$$I_N(A,B) = \frac{2I(A,B)}{H_A + H_B}. \tag{B3}$$

with the obvious range of $0 \leq I_N(A,B) \leq 1$. High I_N values indicate better agreement between compared partitions.

Appendix C: Computational susceptibility

The complexity $\Sigma(e)$ of the energy landscape is related to the number of states $\mathcal{N}(E) \sim \exp[N\Sigma(e)]$ [20] with energy E and energy density $e = E/N$. In the current analysis, we detect the onset of the high complexity with no prior assumptions or approximations by computing a "computational susceptibility" [11] defined as

$$\chi_n = I_N(s = n) - I_N(s = 4). \tag{C1}$$

That is, χ measures the increase in the normalized mutual information I_N as the number of trials (number of independently solved starting points in the energy landscape) $s = n$ is increased. Physically, we evaluate how many different optimization trials are necessary to achieve a desired accuracy threshold.

χ evaluates the expected response of the system to additional optimization effort. That is, a higher χ indicates that additional optimization effort will likely result in a better solution. A low value of χ indicates that there will be less improvement from the additional effort whether due to a trivially solvable system, a complex energy landscape with numerous local minima that trap the solver (at low to moderate temperatures), or thermal-oriented effects of randomly wandering the energy landscape.

Appendix D: Heat Bath Algorithm

We extend the greedy algorithm in [11, 34] to non-zero temperatures by applying a heat bath algorithm. After, we connect the system to a large thermal reservoir at a constant temperature T, the probability for a particular node to move from community a to b is set by a thermal distribution [9],

$$p_{a\to b} = \frac{\exp(-\Delta E_{a\to b}/T)}{\sum_d \exp(-\Delta E_{a\to d}/T)}. \tag{D1}$$

$\Delta E_{a\to b}$ is the energy change that results if the node is moved to the new community b, and the index d runs over

all connected clusters including its current community or a new empty community. The steps of our heat bath algorithm are as follows:

(1) *Initialize the system.* Initialize the network into a "symmetric" state by assigning each node as the lone member of its own community (*i.e.*, $q_0 = N$).

(2) *Find the best cluster for node i.* Select a node and determine to which clusters it is connected (including its current community and an empty cluster). Calculate the energy change $\Delta E_{a\to b}$ required to move to each connected cluster b. Calculate and sum all Boltzmann weights. Generate a random number between 0 and 1 and determine into which cluster the node is placed.

(3) *Iterate over all nodes.* Repeat step 2 in sequence for each node.

(4) *Merge clusters.* Allow for the merger of community pairs based on the same Boltzmann-weighted merge probabilities.

(5) *Repeat the above two steps.* Repeat steps 2 through 4 until the maximum number of iterations is reached.

(6) *Repeat all the above steps for s trials.* Repeat steps 1–5 for s trials and select the lowest energy trial as the best solution. Each trial randomly permutes the order of nodes in the initial state.

This HBA is similar to our greedy algorithm except that we use a random process to select the node moves in steps (2) and (4). The results obtained at low temperature by our HBA are very close to the results obtained by the zero temperature greedy algorithms. Note that there is no cooling scheme as occurs in SA, so step 5 ends at a maximum number of iterations as opposed to a unchanged best partition that is achieved as $T \to 0$ in SA.

In the easy phase, different starting trajectories, each beginning in the symmetric initial state, but they often lead to the same solution. In the hard phase, changing the random seed may significantly alter the final result of an individual trial because the solver becomes trapped in different local minima. Thus we apply additional trials in order to sample different regions of the energy landscape and arrive at better solutions. In the unsolvable phase, increasing the number of trials s does not substantially change the quality of the solutions unless one happens to sample the energy landscape in the immediate vicinity near the optimal partition, but the probability of doing so is small with a finite number of trials s.

Appendix E: Tutte polynomials

We give a very brief introduction to Tutte polynomials consisting of the essential facts necessary for the derivations presented in this paper. The notation used here is mostly standard, but the notation elsewhere in the text deviates from standard notation in order to facilitate the partition function derivation in Sec. VII D. For an undirected graph G, we denote the *deletion* (removal) of an edge e by G' and a *contraction* of the edge by G'' where a contraction consists of removing the edge e and merging the corresponding vertices.

1. Unweighted graph G

If G has no edges, the Tutte polynomial is $t(G; x, y) = 1$. If G is a disjoint graph of partitions, then A and B $t(G; x, y) = t(A; x, y)\, t(B; x, y)$. When an edge e in an unweighted graph G is "*cut*," the recurrence relations are [66]:

- For a general edge, $t(G; x, y) = t(G'_e; x, y) + t(G''_e; x, y)$ which is the sum of two graphs where e is deleted and contracted.
- If edge e is an isthmus between two otherwise disconnected regions of G, then $t(G; x, y) = x\, t(G''_e; x, y)$ where the edge e is contracted.
- If edge e is a loop (a vertex self-edge), then $t(G; x, y) = y t(G'_e; x, y)$ where the edge e is deleted.

The resulting Tutte polynomial is a function of two variables (x, y), and it is independent of the construction order. Different graphs G and H may be described by the same function $t(G; x, y) = t(H; x, y)$. A sample calculation is performed Appendix F for a circle of complete sub-graphs (cliques) as shown in Fig. 23(b).

Tutte polynomials are related to the partition function of a ferromagnetic ($J > 0$) or antiferromagnetic ($J < 0$) Potts model given by

$$H(\{\sigma\}) = -\sum_{i\neq j} J\delta(\sigma_i, \sigma_j) \tag{E1}$$

for any connected pair of nodes i and j with an interaction strength J. The corresponding partition function is

$$Z = q^{k(G)} v^{|V|-k(G)} t(G; x, y) \tag{E2}$$

where q is the number of clusters or states, $v = \exp(\beta J) - 1$, G denotes the graph, $k(G)$ is the number of connected components in G, $|V|$ is the number of vertices, $x = (q + v)/v$ and $y = v + 1$.

In Sec. VII D, we use the following lemma to derive high temperature T approximation for a constructed random graph. We denote K_l as the Tutte polynomial for a complete graph, and $K_l^{(d)}$ denotes that the graph has d duplicated (possibly redundant) edges.

Lemma 1. *For a clique $K_l^{(d)}$ of size l with d duplicate edges between any pair of nodes, the Tutte polynomial at $y = 0$ is K_l.*

Proof. Let G be a complete graph with l vertices and $d = 1$ redundant edge. If we delete and contract the duplicate edge, the Tutte polynomial $t(G) \equiv K_l^{(d=1)}$ will be

$$K_l^{(1)} = K_l + K_{l-1}^{(l-1)}.$$

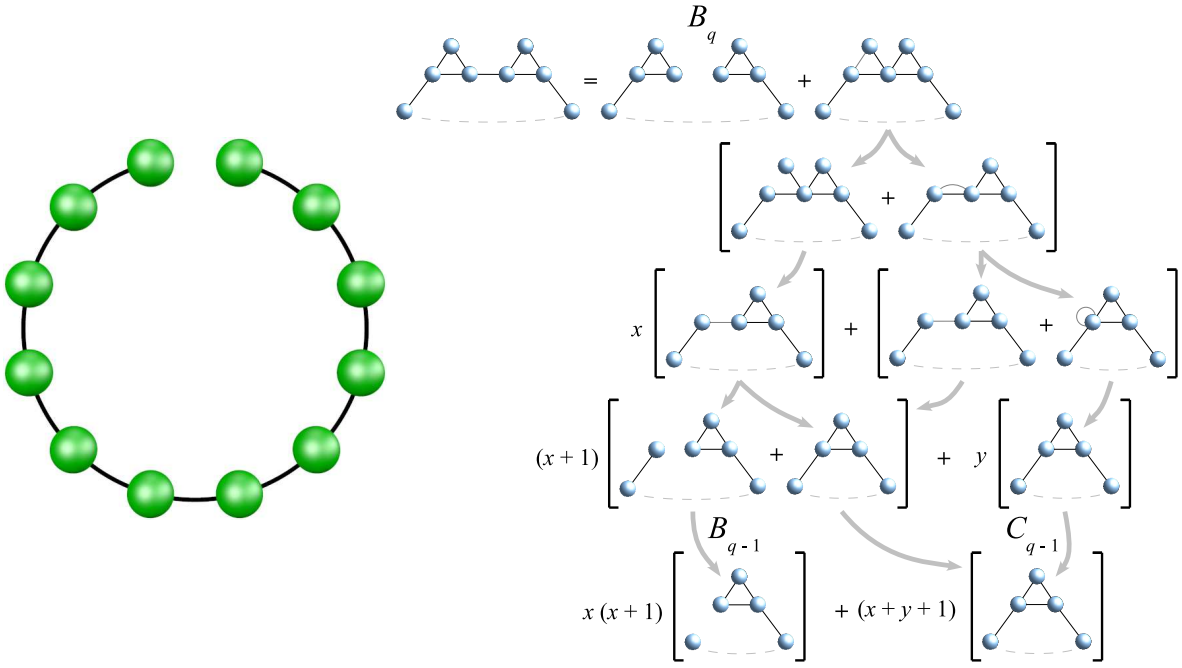


FIG. 23: (Color online) In panel (a), we depict a chain B_q of q cliques (complete sub-graphs of maximally connected clusters) of size l connected by single edges. The corresponding circle of q cliques C_q is depicted in Fig. 21. In panel (b), we show the derivation of the Tutte polynomial in Eq. (F2) for size $l = 3$ cliques. We iteratively break edges and merge nodes according to the Tutte polynomial recurrence relation [66] in Appendix E 1 until we arrive at configurations that are reduced clique circle components. For presentation purposes, gray edges are cut in the next line of the derivation. The dashed gray line at the bottom of each sub-graph represents the remainder of the clique circle which is not touched or affected by the operations on the displayed subgraph.

The contracted vertex in the second term contains $r = 1$ loop. We cut the loop and have

$$K_l^{(1)} = K_l + yK_{l-1}^{(l-2)}|_{y=0} = K_l, \tag{E3}$$

where we invoked the above $y = 0$ condition of the Lemma in the second equality.

Now, assume that we can reduce $K_l^{(d)} = K_l$. Let G be a complete graph with l vertices and $d + 1$ duplicate edges. If we cut one duplicate edge, the resulting Tutte polynomial $t(G) \equiv K_l^{(d+1)}$ will be

$$K_l^{(d+1)} = K_l^{(d)} + K_{l-1}^{(d+l-1)}.$$

The contracted vertex in the second term contains $r \geq 1$ loops. We cut each loop in sequence and obtain

$$K_l^{(d+1)} = K_l^{(d)} + y^r K_{l-1}^{(d+l-r-1)}|_{y=0} = K_l^{(d)}. \tag{E4}$$

If $K_l^{(d)} = K_l$ then by Eq. (E4), we may equate $K_l^{(d+1)} = K_l$. Equation (E3) demonstrates that this relation holds for $d = 1$. Thus, putting all of the pieces together, by mathematical induction $K_l^{(d)} = K_l$ holds true for any integer $d \geq 1$. □

2. Weighted graph G

An excellent summary of multivariate Tutte polynomials (MVTP) is found in Ref. [69]. The MVTP allows for arbitrary weights $\mathbf{v} = [v_e]$ for the edges $\{e\}$ of G. If G has no edges, the MVTP is $Z(G; q, \mathbf{v}) = q$. For an undirected graph G, the weighted Potts Hamiltonian is

$$H(\{\sigma\}) = -\sum_{i \neq j} J_{ij}\delta(\sigma_i, \sigma_j). \tag{E5}$$

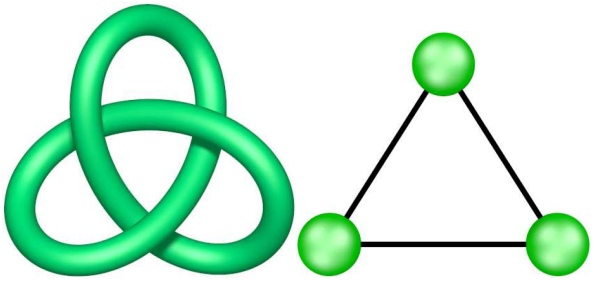
When an edge e in G is "*cut*," the recurrence relation is

$$Z(G; q, \mathbf{v}) = Z(G'; q, \mathbf{v}) + v_e Z(G''; q, \mathbf{v}) \tag{E6}$$

where J_e corresponds to the edge weight between two nodes i and j and $v_e = \exp(\beta J_e) - 1$.

As with the unweighted case, if G is a disjoint graph of partitions A and B, then $Z(G; x, y) = Z(A; q, \mathbf{v})\ Z(B; q, \mathbf{v})$. If partitions A and B are joined at a single vertex, then then $Z(G; x, y) = Z(A; q, \mathbf{v})\ Z(B; q, \mathbf{v})/q$. Unlike Eq. (E2) for unweighted graphs, Eq. (E6) holds for loops or bridges, but for concreteness, cutting an isthmus e yields

$$Z(G; q, \mathbf{v}) = (1 + v_e/q)\, Z(G'_e; x, y) \tag{E7}$$
$$Z(G; q, \mathbf{v}) = (q + v_e)\, Z(G''_e; x, y) \tag{E8}$$



(a) (b)

FIG. 24: (Color online) Panel (a) depicts the trefoil knot, and panel (b) shows the corresponding graph G constructed from the distinct knot regions and crossings [72]. That is, nodes correspond to "checkerboard-shaded" regions (shade the outside lobes of the trefoil knot leaving the interior region unshaded), and edges correspond to knot crossings. Jones polynomials $V_J(x)$ in knot theory are related to Tutte Polynomials, and Eq. (G1) represents the trefoil knot corresponding to the triangle subgraph in panel (b).

where e is deleted or contracted, respectively. If e is a loop, then

$$Z(G;q,\mathbf{v}) = (1+v_e)Z(G'_e;x,y). \tag{E9}$$

Note that the MVTP *is* the partition function. That is, there are no prefactors of q or v_e. Finally, if two parallel edges connect the same pair of nodes i and j with weights J_1 and J_2, then Z_G is unchanged if we replace the parallel edges by a single edge with a weight $J' = J_1 + J_2$ (this negates the need for lemma 1 above).

Appendix F: Derivation of the Tutte polynimial for a circle of cliques

As depicted in Fig. 21, we define C_q as a circle of q cliques where we focus those of size $l = 3$ for the current derivation. The Tutte polynomial for a triangle is $\Delta \equiv \left(x^2+x+y\right)$. For convenience, we also define, $\Delta' \equiv (\Delta + x + 1) = [(x+1)^2 + y]$ and $y' \equiv (x+y+1)$.

We define B_q to be the Tutte polynomial for a clique chain as depicted in Fig. 23(a). In this case, it is trivial to construct B_q

$$B_q = x^{q-1}\left(x^2+x+y\right)^q. \tag{F1}$$

With Eq. (F1), we construct a recurrence relation for the clique circle configurations as shown in Fig. 23(b)

$$C_q = B_q + x\,(x+1)\,B_{q-1} + (x+y+1)\,C_{q-1}. \tag{F2}$$

From this relation, we can sum the series exactly.

$$\begin{aligned} C_q &= B_q + \Delta' B_{q-1} + x\,(x+1)\,(x+y+1)\,B_{q-2} \\ &\quad + \,(x+y+1)^2\,C_{q-2} \\ &\vdots \qquad\qquad\qquad \vdots \\ C_q &= B_q + \Delta' \sum_{i=0}^{q-4} (x+y+1)^i\,B_{q-i-1} \\ &\quad + \Delta\,(x+y+1)^{q-3}\,B_2 + (x+y+1)^{q-2}\,C_2. \end{aligned} \tag{F3}$$

Note that the last B_j term uses Δ not Δ'. Also, it can be shown that $C_2 = (x+1)^2\left(x^3+\Delta\right) + y\,(x+1)\,\Delta$. Substituting these values into the equation, we arrive at

$$\begin{aligned} C_q &= x^{q-1}\Delta^q + \Delta' \sum_{i=0}^{q-4} y'^i x^{q-i-2}\Delta^{q-i-1} + xy'^{q-3}\Delta^3 \\ &\quad + xy'^q\left(x^2+x+1\right) + yy'^{q-1}\Delta'. \end{aligned} \tag{F4}$$

In the high temperature T limit, $y \ll x$, so we approximate $y \simeq 0$, and the equation simplifies to

$$\begin{aligned} C_q^{(T)} &\simeq x\,(x+1)^q\left[x^{2q-2}+\cdots+x^2+x+1\right] \\ &= x\,(x+1)^q\left[\frac{1-x^{2q-1}}{1-x}\right], \end{aligned} \tag{F5}$$

We make a final high T approximation

$$C_q^{(T)} \simeq (x+1)^{q+1}\,x^{2q-3} \tag{F6}$$

using $\left(x^{2q-1}-1\right) \simeq x^{2q-1}$ and $(1-x)^{-1} \simeq (1+x)/x^2$.

Appendix G: Random knot "transitions"

A general 3D knot may be represented as a 4-valent planar graph [72] [*i.e.*, corresponding to a two-dimensional (2D) square lattice connectivity allowing self-loops]. This relation connects the Tutte polynomial to the Jones polynomial in knot theory. Conversely, all connected, signed planar graphs have a corresponding link diagram representation (2D knot projection). Alternating over-under crossings result in unsigned planar graphs [72] (e.g., the trefoil knot in Fig. 24). Ref. [73] provides an introduction to the mathematics and physics of knot theory. The Jones polynomial of a given knot is intimately related to quantum field theories [74], via its connection to [an SU(2) type] Wilson loop associated the same knot.

As a concrete example, Fig. 24(a) depicts a simple trefoil knot which is related to the triangle clique depicted in Fig. 24(b) [72]. The Tutte polynomial of Fig. 24(b) is $K_3(G;x,y) = x^2+x+y$. Then we generate the Jones polynomial

$$V_J(x) = x^2 + x + \frac{1}{x} \tag{G1}$$

where we used $xy = 1$ because the trefoil knot has alternating crossings [75]. While the trefoil knot is clearly not random, we conjecture that the transitions detected in random graphs with embedded ground states in the current work can have similar transition repercussions in random knots.

[1] M. E. J. Newman, Phys. Today **61**, 33 (2008).
[2] S. Fortunato, Phys. Rep. **486**, 75 (2010).
[3] F. Radicchi,C. Castellano,F. Cecconi,V. LoretoandD. Parisi Proc. Natl. Acad. Sci. U.S.A **101**, 2658 (2004).
[4] A. Lancichinetti, F. Radicchi and J Ramasco, Phys. Rev. E **81**, 046110 (2010).
[5] F. Radicchi, A. Lancichinettiand J Ramasco, Phys. Rev. E **82**, 026102 (2010).
[6] M. Rosvall and C. T. Bergstrom, Proc. Natl. Acad. Sci. U.S.A. **105**, 1118 (2008).
[7] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, J. Stat. Mech.: Theory Exp. **10**, P10008 (2008).
[8] M. B. Hastings, Phys. Rev. E **74**, 035102(R) (2006).
[9] J. Reichardt and S. Bornholdt, Phys. Rev. E **74**, 016110 (2006).
[10] A. Lancichinetti, S. Fortunato, and J. Kertész, New J. Phys. **11**, 033015 (2009).
[11] P. Ronhovde and Z. Nussinov, Phys. Rev. E **80**, 016109 (2009).
[12] X.-Q. Cheng and H.-W. Shen, J. Stat. Mech.: Theory Exp. **2010**, P04024 (2010).
[13] V. Gudkov, V. Montealegre, S. Nussinov, and Z. Nussinov, Phys. Rev. E **78**, 016113 (2008).
[14] M. J. Barber and J. W. Clark, Phys. Rev. E **80**, 026129 (2009).
[15] L. Danon, A. Díaz-Guilera, J. Duch, and A. Arenas, J. Stat. Mech.: Theory Exp. **9**, P09008 (2005).
[16] A. Noack and R. Rotta, in *Experimental Algorithms*, edited by J. Vahrenhold (Springer-Verlag Berlin, Heidelberg, 2009), vol. 5526, pp. 257–268.
[17] H.-W. Shen and X.-Q. Cheng, J. Stat. Mech.: Theory Exp. **2010**, P10020 (2010).
[18] A. Lancichinetti and S. Fortunato, Phys. Rev. E **80**, 056117 (2009).
[19] T. Hogg, B. A. Huberman, and C. P. Williams, Artificial Intelligence **81**, 1 (1996).
[20] M. Mézard, G. Parisi, and R. Zecchina, Science **297**, 812 (2002).
[21] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Nature (London) **400**, 133 (1999).
[22] S. Mertens, Phys. Rev. Lett. **81**, 4281 (1998).
[23] I. P. Gent and T. Walsh, Artificial Intelligence **88**, 349 (1996).
[24] M. Weigt and A. K. Hartmann, Phys. Rev. Lett. **84**, 6118 (2000).
[25] L. Lacasa, B. Luque, and O. Miramontes, New J. Phys. **10**, 023009 (2008).
[26] F. Krzakala and L. Zdeborová, Journal of Physics: Conference Series **95**, 012012 (2008).
[27] H. Bauke, S. Mertens, and A. Engel, Phys. Rev. Lett. **90**, 158701 (2003).
[28] G. Mukherjee and S. S. Manna, Phys. Rev. E **71**, 066108 (2005).
[29] R. Arévalo, I. Zuriguel, and D. Maza, Phys. Rev. E **81**, 041302 (2010).
[30] B. Ashok and T. K. Patra, Pramana **75**, 549 (2010).
[31] K. Rose, E. Gurewitz, and G. C. Fox, Phys. Rev. Lett. **65**, 945 (1990).
[32] T. Graepel, M. Burger, and K. Obermayer, Phys. Rev. E **56**, 3876 (1997).
[33] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).
[34] P. Ronhovde and Z. Nussinov, Phys. Rev. E **81**, 046114 (2010).
[35] D. Hu, P. Ronhovde, and Z. Nussinov, Phil. Mag. **92**, 406 (2012).
[36] B. H. Good, Y.-A. de Montjoye, and A. Clauset, Phys. Rev. E **81**, 046106 (2010).
[37] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, Phys. Rev. Lett. **107**, 065701 (2011).
[38] E. J. Lee, K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. E **71**, 056108 (2005).
[39] Y. Moreno, R. Pastor-Satorras, A. Vzquez, and A. Vespignani, Europhys. Lett. **62**, 292 (2003).
[40] W.-X. Wang and G. Chen, Phys. Rev. E **77**, 026101 (2008).
[41] J.-F. Zheng, Z.-Y. Gao, and X.-M. Zhao, Europhys. Lett. **79**, 58002 (2007).
[42] J.-j. Wu, Z.-y. Gao, and H.-j. Sun, Phys. Rev. E **74**, 066111 (2006).
[43] Y. Ikeda, T. Hasegawa, and K. Nemoto, Journal of Physics: Conference Series **221**, 012005 (2010).
[44] A. Tahbaz-Salehi and A. Jadbabaie, in *Proceedings of the 2007 American Control Conference* (IEEE, 2007), pp. 699–704.
[45] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).
[46] M. Á. Serrano and M. Boguñá, Phys. Rev. Lett. **97**, 088701 (2006).
[47] C. Moore and M. E. J. Newman, Phys. Rev. E **61**, 5678 (2000).
[48] A. Lancichinetti, S. Fortunato, and F. Radicchi, Phys. Rev. E **78**, 046110 (2008).
[49] V. A. Traag, P. Van Dooren, and Y. Nesterov, Phys. Rev. E **84**, 016114 (2011).
[50] M. E. J. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).
[51] S. Fortunato and M. Barthélemy, Proc. Natl. Aca. Sci. U.S.A. **104**, 36 (2007).
[52] A. Lancichinetti and S. Fortunato, Phys. Rev. E **84**, 066122 (2011).
[53] P. W. Kasteleyn and C. M. Fortuin, in *Proceedings of the International Conference on Statistical Mechanics, September 9-14, Koyto* (1969), vol. 26, p. 11.
[54] U. N. Raghavan, R. Albert, and S. Kumara, Phys. Rev. E **76**, 036106 (2007).
[55] M. T. Mercaldo, J.-C. Anglès d'Auriac, and F. Iglói, Phys. Rev. E **69**, 056112 (2004).
[56] J.-C. A. d'Auriac, F. Iglói, M. Preissmann, and A. Sebö, J. Phys. A **35**, 6973 (2002).

[57] R. Juhász, H. Rieger, and F. Iglói, Phys. Rev. E **64**, 056122 (2001).
[58] P. Ronhovde and Z. Nussinov, (in preparation) (2012).
[59] P. Ronhovde, D. Hu, and Z. Nussinov, Europhys. Lett. **99**, 38002 (2012).
[60] D. Hu, P. Ronhovde, and Z. Nussinov, Phys. Rev. E **85**, 016101 (2012).
[61] R. Juhász, H. Rieger, and F. Iglói, Phys. Rev. E **64**, 056122 (2001).
[62] M. T. Mercaldo, J.-C. A. d'Auriac, and F. Iglói, Europhys. Lett. **70**, 733 (2005).
[63] S.-C. Chang and R. Shrock, Int. J. Mod. Phys. B **21**, 979 (2007).
[64] L. Mittag and M. J. Stephen, J. Phys. A: Math. Nucl. Gen. **7**, L109 (1974).
[65] P. A. Pearce and R. B. Griffiths, J. Phys. A: Math. Gen. **13**, 2143 (1980).
[66] D. J. A. Welsh and C. Merino, J. Math. Phys. **41**, 1127 (2000).
[67] V. A. Traag and J. Bruggeman, Phys. Rev. E **80**, 036115 (2009).
[68] J. Reichardt and S. Bornholdt, Phys. Rev. Lett. **93**, 218701 (2004).
[69] B. Jackson and A. D. Sokal, J. Combinatorial Theory, Series B **99**, 869 (2009).
[70] C. D. Batista and Z. Nussinov, Phys. Rev. B **72**, 045137 (2005).
[71] Z. Nussinov, G. Ortiz, and E. Cobanera, e-print arXiv:1110.2179 (2011).
[72] L. H. Kauffman, Discrete Appl. Math **25**, 105 (1989).
[73] L. H. Kauffman, Rep. on Progress in Phys. **68**, 2829 (2005).
[74] E. Witten, Comm. Math. Phys. **121**, 351 (1989).
[75] M. B. Thistlethwaite, Topology **26**, 297 (1987).