Mean-field approximation for structural balance dynamics in heat bath

Krzysztof Malarz [1,*] and Janusz A. Hołyst [2,†]

[1]*AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, al. Mickiewicza 30, 30-059 Kraków, Poland*

[2]*Faculty of Physics, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland*

(Received 24 June 2022; accepted 14 December 2022; published 30 December 2022)

A critical temperature for a complete signed graph of N agents where the time-dependent polarization of links tends towards the Heider (structural) balance is found analytically using the heat-bath approach and the mean-field approximation as $T^c = (N-2)/a^c$, where $a^c \approx 1.716\,49$. The result is in perfect agreement with numerical simulations starting from the paradise state where all links are positively polarized as well as with the estimation of this temperature received earlier with much more sophisticated methods. When heating the system, one observes a discontinuous and irreversible phase transition at T^c from a nearly balanced state when the mean link polarization is about $x_c = 0.796\,388$ to a disordered and unbalanced state where the polarization vanishes. When the initial conditions for the polarization of links are random, then at low temperatures a balanced bipolar state of two mutually hostile cliques exists that decays towards the disorder and there is a discontinuous phase transition at a temperature T^d that is lower than T^c. The system phase diagram corresponds to the so-called fold catastrophe when a stable solution of the mean-field equation collides with a separatrix, and as a result a hysteresislike loop is observed.

DOI: 10.1103/PhysRevE.106.064139

I. INTRODUCTION

Let i and j be persons and k be an object that could be a third person, article, idea, event, etc. If the persons i and j possess the same attitude towards the object k (e.g., they both like or both dislike it), then the theory of structural balance postulated by Heider [1] says that it is more probable there is a positive relation between i and j. On the other hand, if there is a disagreement between the attitudes of i and j towards k, then it is more likely that there is a negative relation between i and j. In the case where k is a person, the above propositions can be formed as the following rules: a friend of my friend is my friend, an enemy of my enemy is my friend, a friend of my enemy is my enemy, and an enemy of friend is my enemy.

Structural balance theory met a lot of interest in social science and it was observed in many social groups when friendships and antipathies could be detected (see, e.g., Refs. [2–12]). Antal, Krapivsky, and Redner [13,14] found a way to use the master equation for a description of possible dynamics of complex networks evolving toward the structural balance when mutual attitudes are described by binary variables of corresponding links.

Nowadays there are several attempts at theoretical description and computational simulation of the Heider balance, for example, when the attitudes are continuously changing variables [15–17] or when attitudes are link attributes following from Hamming distances between nodes attributes [18]. Often, an interpersonal relationship between two agents evolves, driven by the products of their relations (positive or negative) with their common neighbors. In this way, two friendly or two hostile relations of two agents with their common neighbor improve their mutual relation, while their different relations with a neighbor drive their mutual relation to hostility. For example, in Refs. [15,16], the relations are represented by real numbers, and the system dynamics is represented by a set of differential equations. In these papers, the network of relations is a complete graph. In Refs. [19,20] the relations are either positive or negative, and the dynamics is defined by a cellular automaton deterministic or with a thermal noise and a local neighborhood of different range. In Refs. [21,22] the relations are also discrete, the automaton rule is deterministic and the topology is a complete graph. In all these approaches, the target is a balanced state, i.e., a partition of the graph into two mutually hostile but internally friendly groups. When the sign of the relation of a pair as in Ref. [22] is assumed to oscillate with the sum of products of relations over their neighbors, this target is reached immediately—in one time step—for each initial state [23]. For other approaches see Ref. [24].

Recently [25–31], the Heider's dynamics has been enriched with social temperature T [32]. In Ref. [25] authors show that in the investigated system the first-order phase transition from an ordered to a disordered state is observed. A system of two coupled algebraic equations has been received using a mean-field approximation for an average link polarization and a correlation function between neighboring links' polarization. The critical temperature of the complete graph consisting of 50 nodes has been estimated by numerical solutions of these equations and agent-based simulations have confirmed the presence of the phase transition in such a model. In this paper, we show a much simpler theoretical approach leading to the same conclusions. Similarly as in Ref. [25] we use the mean-field approximation, but we

*malarz@agh.edu.pl

†janusz.holyst@pw.edu.pl

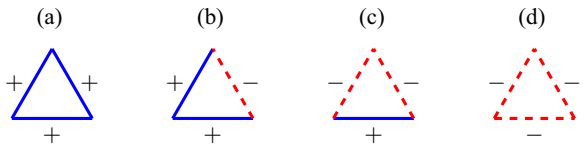

FIG. 1. Heider's triads corresponding to balanced (first and third) and imbalanced (second and fourth) states

find the critical temperature T_c only from the average value of link polarization and we show that T_c is proportional to the number M of different triangles containing a given link $T^c = M/1.716\,49\cdots$. If a network with all positive links is heated, then the discontinuous phase transition takes place when the mean link polarization decays to the critical value $x_c = 0.796\,388$. Our analytical results are well supported by computer simulations.

II. MODEL

Consider a network of N agents, and let us assume that the polarization of the links between two agents i and j is $x_{ij} = \pm 1$. The dynamics towards the Heider balance can be written as

$$x_{ij}(t+1) = \mathrm{sgn}\left(\sum_k^{M_{ij}} x_{ik}(t)x_{kj}(t)\right), \tag{1}$$

where the summation goes through M_{ij} common nearest neighbors of the connected nodes i and j; that is, M_{ij} is the number of triangles that involve the link ij. It means that for a single triangle system presented in Fig. 1, the first and the third triangles are balanced in the Heider's sense (as a friend of my friend is my friend and an enemy of my friend is my enemy), while the second and the fourth triangles are not. In the latter case actors at triangles nodes either encounter the cognitive dissonance—as they cannot imagine how his/her friends can be enemies—or everybody hates everybody, which should lead to the creation of a two-against-one coalition. Let us stress that formally the sum $\sum_k^{M_{ij}} x_{ik}(t)x_{kj}(t)$ can be treated as a local field acting on the link x_{ij} and this field follows from the Hamiltonian [13,25]

$$\mathcal{H} = -\sum_{i>j>k}^{N} x_{ij}x_{jk}x_{ki}. \tag{2}$$

III. MEAN-FIELD ANALYTICAL APPROACH

If one assumes that the link dynamics possesses a probabilistic character, then a natural form of updating rule (1) as in Heider dynamics can be the following:

$$x_{ij}(t+1) = \begin{cases} +1 & \text{with probability } p_{ij}(t), \\ -1 & \text{with probability } 1 - p_{ij}(t), \end{cases} \tag{3a}$$

where

$$p_{ij}(t) = \frac{\exp(\xi_{ij}(t)/T)}{\exp[\xi_{ij}(t)/T] + \exp[-\xi_{ij}(t)/T]} \tag{3b}$$

and

$$\xi_{ij}(t) = \sum_k^{M_{ij}} x_{ik}(t)x_{kj}(t). \tag{3c}$$

Here, the positive variable T can be considered as a social temperature [32] (or a measure of the noise amplitude) and in the limit $T \to 0^+$ we have $p \to 1$ so Eq. (3) reduces to Eq. (1). Equation (3) is nothing else but the heat-bath algorithm [33, p. 505] for a stochastic version of Eq. (1), and thus it is ready for direct implementation in analytical investigations and computer simulations.

The expected value $\langle x_{ij}(t+1)\rangle$ in this approach is equal to

$$\langle x_{ij}(t+1)\rangle = \tanh\left(T^{-1}\sum_k^{M_{ij}} x_{ik}(t)x_{kj}(t)\right), \tag{4}$$

where $\langle\cdots\rangle$ stands for a mean value related to the stochastic process defined by Eq. (3).

The mean $\langle x_{ij}(t+1)\rangle$ is a continuous variable that can be negative or positive, and for $T \to 0^+$, Eq. (4) reduces to Eq. (1).

Now in our mean-field approximation we write the correlation function $\langle x_{ik}(t)x_{kj}(t)\rangle$ as the product $\langle x_{ik}(t)\rangle\langle x_{kj}(t)\rangle$. Let us note that such an approximation for the correlation function of link polarization is similar to the mean-field assumption used for the Ising model where correlations between spins S_i and S_j are neglected, i.e., $\langle S_iS_j\rangle = \langle S_i\rangle\langle S_j\rangle$ (see, e.g., Ref. [34]). In fact, in Ref. [25] another mean-field approach was proposed where link-link correlations were considered but as we demonstrate in the Appendix they can be disregarded in the thermodynamical limit of the studied system. Then instead of Eq. (4) we have

$$\{\langle x_{ij}(t+1)\rangle\} \approx \tanh\left(\frac{1}{T}\sum_k^{M_{ij}}\{\langle x_{ik}(t)\rangle\}\{\langle x_{kj}(t)\rangle\}\right), \tag{5}$$

where $\{\cdots\}$ denotes the average over all $N(N-1)/2$ available nodes' pairs.

Now let us assume—also in agreement with the spirit of mean-field approximation—that all averages are the same:

$$\{\langle x_{ij}\rangle\} = \{\langle x_{ik}\rangle\} = \{\langle x_{kj}\rangle\} = x. \tag{6}$$

The above approximations are justified in the neighborhood of the paradise[1] state where the majority of triangles of the type shown in Fig. 1(a) are present. However, as we show in further numerical simulations, the approach also works well in states far from paradise ($x \approx 0$).

It follows that we get

$$x(t+1) = \tanh[ax^2(t)], \tag{7}$$

where

$$a = M/T, \tag{8}$$

and $M = \{M_{ij}\}$ is the average number of common neighbors of agents i and j (M also indicates the number of different triads containing the edge ij).

[1]For paradise states, all relations are friendly [13].

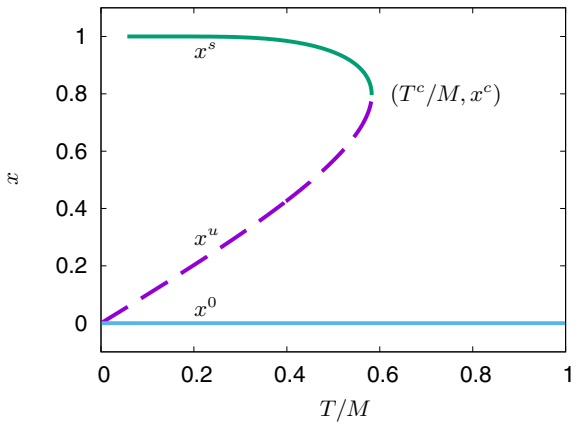

FIG. 2. Solutions for mean values of links polarizations in Heider balance resulting from the mean-field approximation $x = \tanh(ax^2)$. At the point $T^c/M = 1/a^c \approx 0.582\,584\cdots$ $(x^c \approx 0.796\,388\cdots)$, a discontinuous transition between the upper branch and the solution $x = 0$ takes place.

We immediately recognize $x^0 = 0$ as a stable fixed point for any value of the a parameter, and for $a \ll 1$ this is the only fixed point of Eq. (7). However, for $a \gg 1$ there are two other fixed points, $x^u < x^s$ corresponding to unstable $x^u(a)$ and stable $x^s(a)$ solutions. In fact, x^u is a separatrix between the domain of attractions of fixed points x^0 and x^s. When $a \gg 1$ then $x^u \approx 1/a$. When the parameter a decreases from high values (this means that the temperature T increases), then the fixed points x^u and x^s coincide together with the point x^c for a certain value of $a = a^c$ (see Fig. 2).

This means that for $a > a^c$ the system is bistable and for $a < a^c$ the system is mono-stable. The above values a^c and x^c can be received from a pair of transcendental algebraic relations that describe the fixed point and its tangency condition, namely,

$$x^c = \tanh[a^c(x^c)^2] \tag{9a}$$

and

$$2a^c x^c = \cosh^2[a^c(x^c)^2]. \tag{9b}$$

The solutions (see Fig. 2) are

$$x^c_{\rm th} \approx 0.796\,388\cdots \tag{10a}$$

and

$$a^c_{\rm th} = \tanh^{-1}(x^c_{\rm th})/\left(x^c_{\rm th}\right)^2 \approx 1.716\,49\cdots. \tag{10b}$$

Let us note that since $x^c_{\rm th} > 0$ a system can express the phenomenon of hysteresis. It also means that we should not observe the values $0 < x < x^c_{\rm th}$ as stable solutions.

IV. NUMERICAL ESTIMATION OF SYSTEM CRITICAL TEMPERATURE

To verify the analytical results in a computer simulation, we directly apply Eq. (3) to the time evolution of x_{ij} for the complete graph with N nodes. For the complete graph the average number of pair neighbors of nodes ij is equal to $M = \{M_{ij}\} = N - 2$, and thus according to Eq. (8) one should expect

$$a^c_{\rm nu} = (N-2)/T^c_{\rm nu}. \tag{11a}$$

To find the value of $T^c_{\rm nu}$ we start the simulation with $T = 0$ and scan the temperature T with step ΔT and look for a value of T^* for which $\langle x_{ij}\rangle$ is positive but for $T^* + \Delta T$ is zero. The true value of $T^c_{\rm nu}$ is hidden somewhere in the interval $[T^*, T^* + \Delta T]$. We assume that the T^c value is uniformly distributed in the interval $[T^*, T^* + \Delta T]$ which allows us to estimate its uncertainty as $u(T^c_{\rm nu}) = \Delta T/\sqrt{3}$. The estimated value of $T^c_{\rm nu} = (T^* + T^* + \Delta T)/2$. Based on Eq. (8) we calculate the value of $a^c_{\rm nu}$ and we can estimate its expanded uncertainty as

$$U\left(a^c_{\rm nu}\right) = k\left|\frac{\partial a}{\partial T}\right|_{T=T^c} u\left(T^c_{\rm nu}\right) = k\frac{N-2}{\left(T^c_{\rm nu}\right)^2}u\left(T^c_{\rm nu}\right), \tag{11b}$$

with the coverage factor $k = 3$ [35].

In Fig. 3 the time evolution of $\{x_{ij}\}$ for various values of social temperature T and various system sizes N is presented. The starting point of the simulation is the homogeneous state (paradise) with $\{x_{ij}\} = +1$ and the scanning temperature step is set to $\Delta T = 1$. The solid red line corresponds to $x^c_{\rm th}$ given by Eq. (10a).

The obtained critical temperatures $T^c_{\rm nu}$ and their uncertainties $U(T^c_{\rm nu})$ are collected in Table I. The obtained values of $a^c_{\rm nu}$ coincide nicely with those obtained analytically [see Eq. (10b)], even under very crude assumptions given by Eq. (7). The values of $a^c_{\rm nu}$ agree within expanded uncertainties $U(a^c_{\rm nu})$ with its analytical partner $a^c_{\rm th}$.

In Figs. 4(a) and 4(c), the dependencies of $\overline{\{x_{ij}\}}$ vs T for $N = 200$ and $N = 50$ are presented. The averaging symbol $\overline{\cdots}$ represents the time average in the last $\tau = 100$ time steps of the simulation, and this time average should be approximately equal to the average $\langle\cdots\rangle$ used in Eq. (5), which comes from the ergodic theorem. Solid symbols correspond to the starting point $\forall i, j : x_{ij} = +1$, while open symbols represent a random initial state $\{x_{ij}\} = 0$. The latter recovers $x^0 = 0$ mentioned earlier.

TABLE I. The numerically obtained values of $T^c_{\rm nu}$ and $a^c_{\rm nu}$ calculated based on Eq. (11a) together with their estimated expanded uncertainties $U(a^c_{\rm nu})$. The uncertainty of $T^c_{\rm nu}$ is $u(T^c_{\rm nu}) = 1/\sqrt{3}$, while $U(a^c_{\rm nu})$ is calculated based on Eq. (11b). Note that the differences between numerically estimated values and analytical results (values in the last column) are smaller than the uncertainties $U(a^c_{\rm nu})$.

N	$T^c_{\rm nu}$	$a^c_{\rm nu}$	$U(a^c_{\rm nu})$	$a^c_{\rm nu} - a^c_{\rm th}$
25	11.5	2.00	0.30	0.28
50	26.5	1.811	0.118	0.095
100	55.5	1.766	0.055	0.049
200	114.5	1.7293	0.0262	0.0128
400	320.5	1.7267	0.0130	0.0102
800	463.5	1.7217	0.0064	0.0052

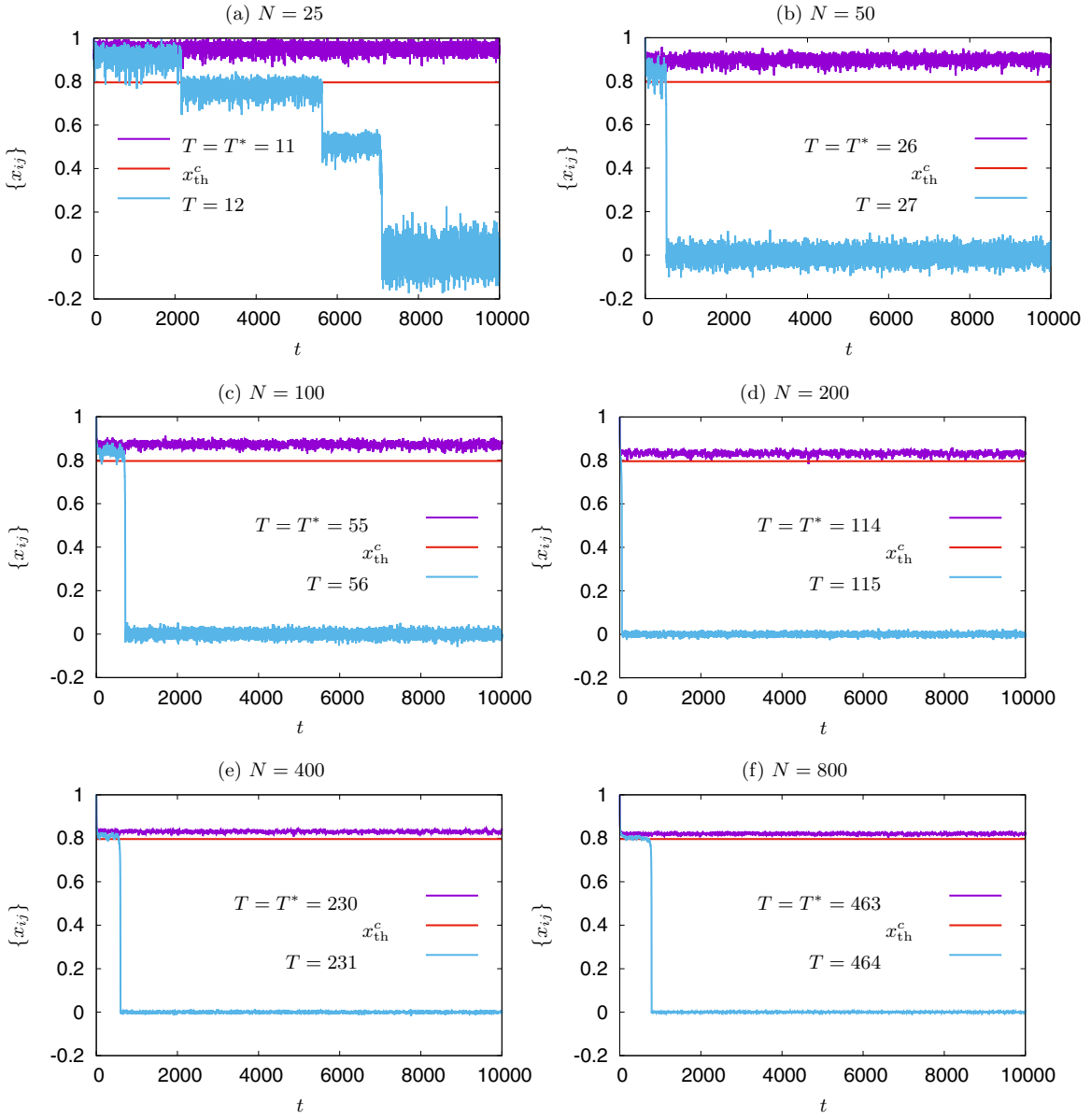

FIG. 3. The time evolution of the average values of all links' polarization $\{x_{ij}\}$ for various social temperatures T and various system sizes N. The starting point of simulation is the homogeneous state (paradise) with $\{x_{ij}\} = +1$ and the scanning temperature step is $\Delta T = 1$. When the temperature $T = T^*$ then averages $\{x_{ij}\}$ oscillate around values that are close but always larger than the critical solution x^c [the solid red line given by the mean-field approach (10a)]; however, when the temperature T is slightly above T^* the system evolves towards the state $\{x_{ij}\} \approx x^0 = 0$.

In Figs. 4(b) and 4(d), dependencies of the system energy density [average value of the Hamiltonian (2) per triangle]

$$E = -\frac{\sum_{i>j>k} x_{ij}x_{jk}x_{ki}}{\binom{N}{3}} \tag{12}$$

are presented. There is a discontinuous change in the mean system energy at the critical temperature; this corresponds to Fig. 4(b) in Ref. [25]. According to the mean-field approximation, Eq. (6), we expect $E = -x^3$, and this approximation is marked by a solid blue line in Figs. 4(b) and 4(d). Similarly to the numerically obtained values of $\{x_{ij}\}$, also values of $\overline{E}$ agree fairly with the proposed mean-field approximation.

V. INFLUENCE OF MUTUALLY HOSTILE CLIQUES ON CRITICAL BEHAVIOR

In previous sections we estimated analytically and numerically the value of the system critical temperature T^c when initial conditions were close to the paradise state. We observed, however, that when a random initial state $\{x_{ij}(t = 0)\} = 0$ was used in our numerical simulations a transition to a phase with a higher energy $\overline{E} = 0$ took place at a temperature T^d that was lower than T^c. Below we discuss the nature of this transition (see also Ref. [27]).

For random initial conditions the observed average $\{x_{ij}\}$ fluctuates around zero in time but the mean energy density is $\overline{E} = -1$ (see Fig. 4) in low temperatures. In other words,

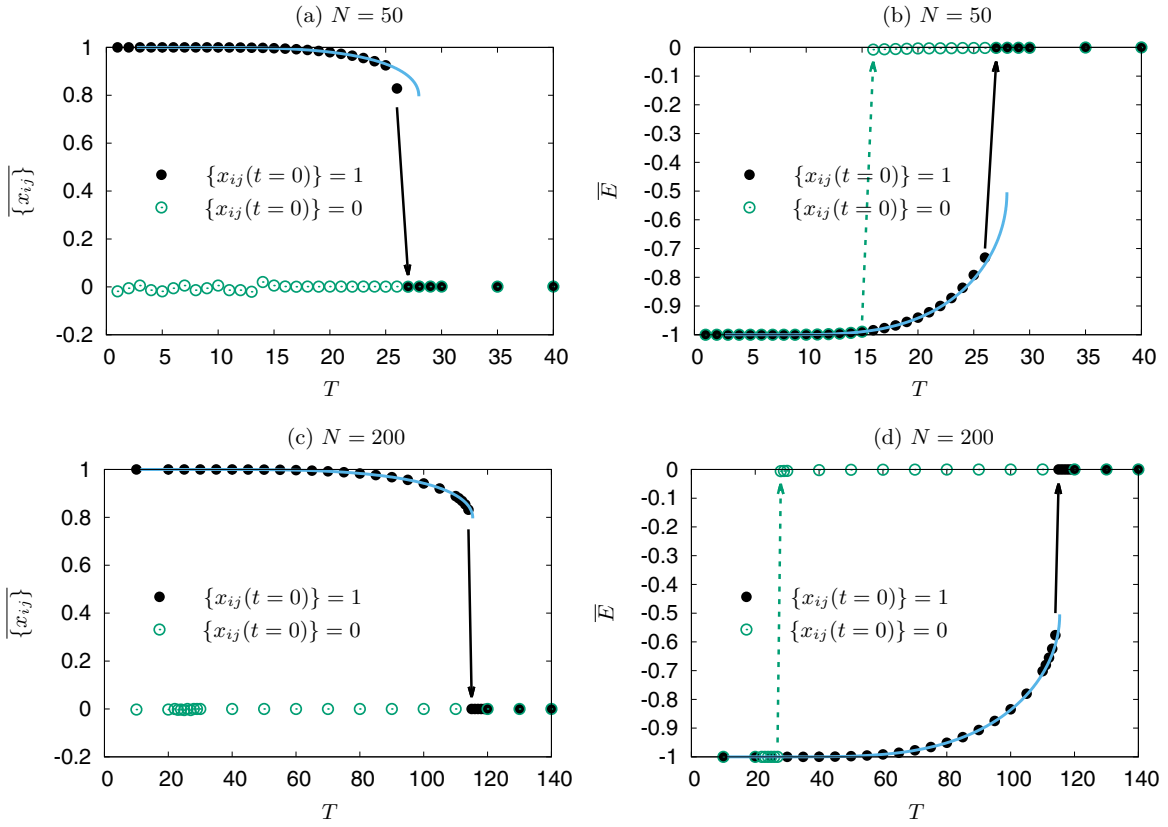

FIG. 4. . Panels (a) and (c) show the averages $\overline{\{x_{ij}\}}$ and panels (b) and (d) show $\overline{E}$ as functions of the temperature T for the complete graphs with $N = 50$ nodes [panels (a) and (b)] and $N = 200$ nodes [panels (c) and (d)]. The averaging symbol $\overline{\cdots}$ stands for time average over the last $\tau = 1000$ time steps of simulation. The solid symbols correspond to the paradise state as the starting point. The open symbols stand for the random initial state $\{x_{ij}\} = 0$. The solid blue lines indicate the mean-field approximation predictions $\overline{\{x_{ij}\}} = x$ [panels (a) and (c)] and $\overline{E} = -x^3$ [panels (b) and (d)]. The solid black or dashed green arrow marks the position of the critical temperature T^c or T^d (see Sec. V).

this energy is the same as the system energy corresponding to the paradise state (with only positive links) $\{x_{ij}\} = 1$. This means that the ground state of the system is degenerated [36]. Although the fact that the state $\{x_{ij}\} = 0$ is stable in time follows from Eq. (7), its nature can be seen to be strange, since naively one would expect a picture of a disordered system corresponding to many unbalanced triangles with $\overline{E} = 0$ and not $\overline{E} = -1$.

The truth is that the stochastic evolution equation, Eq. (3), starting from the unordered state $\{x_{ij}\} = 0$ can lead to a polarized phase consisting of two paradises [cliques containing only friendship triangles of the type shown in Fig. 1(a)] of similar sizes (for entropic reasons) connected by hostile triangles of the type shown in Fig. 1(c)—see Fig. 5. Such a balanced state composition has already been predicted for systems driven by the structural balance by Cartwright and Harary [3] in the middle 1950s. The phase was observed also in the model of constrained triad dynamics (CTD) introduced by Antal *et al.* [13,14] when a randomly selected link flips its temporal polarization x_{ij} only if the total number of imbalanced triad decreases (see also Ref. [19]). Such a case corresponds to Eq. (1) and it is equivalent to the $T \to 0^+$ limit of the heat-bath equations, Eqs. (3).

Let us assume that the size of the entire group is an even number $N = 2h$ and that each hostile group has the size h. Such a partition is the most probable division of the entire group [13,14] and it is then easy to show that the mean of $\{x_{ij}\}$ is zero in such a bipolar phase. In fact, every node possesses $(h-1)$ positive links to agents in the own group and h negative links to the second group. It follows that the mean link polarization $\{x_{ij}\}$ is equal to $-1/(2h-1)$ and vanishes in the thermodynamic limit ($h \to \infty$). Then, the number of triangles of the type shown in Fig. 1(a),

$$n_\Delta(+3) = 2\binom{h}{3} = \frac{1}{3}h(h-1)(h-2), \tag{13}$$

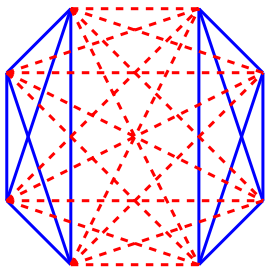

FIG. 5. An example of a bipolar state with two mutually hostile cliques of sizes $h = 4$. Each node possesses $(h-1)$ positive links to agents in the own clique and h negative links to the second clique.

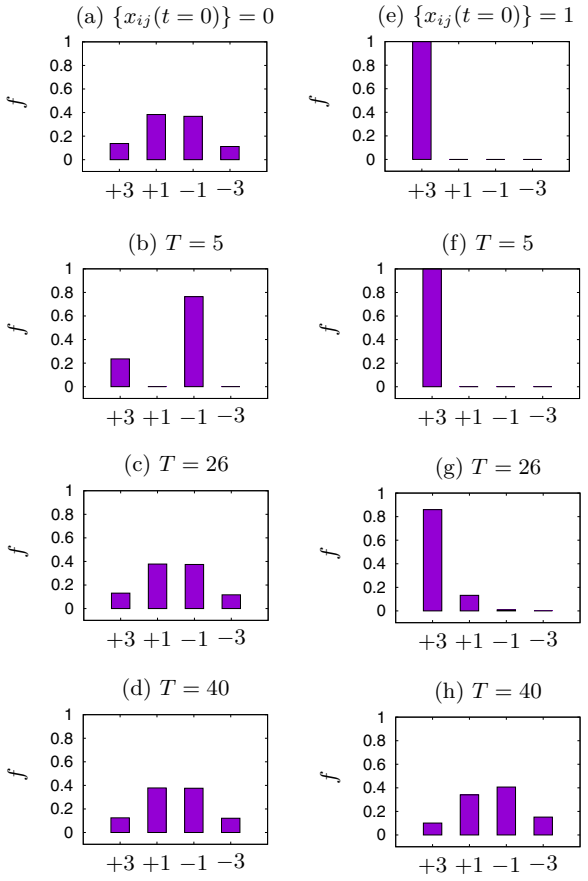

FIG. 6. Frequencies f of the triangles presented in Fig. 1 for $N=50$ and various initial states. In the first row the histograms for initial states are presented. In left (right) column histograms for random (paradise) initial states are presented. In rows 2–4 the histograms of frequencies of various triangles types in the final state of the system are presented. The assumed temperatures T are indicated for each figure part. The numbers on the abscissa axis are the sum of polarizations x_{ij} in a given triangle presented in Fig. 1.

and the number of triangles of the type shown in Fig. 1(c),

$$n_\Delta(-1)=2h\binom{h}{2}=h^2(h-1), \tag{14}$$

may be calculated with combinatorial analysis due to the complete graph symmetry. It follows that, in the thermodynamic limit, there is a special ratio between the numbers of triangles of the type shown in Fig. 1(c) and the number of triangles of the type shown in Fig. 1(a) in this phase,

$$\lim_{h\to\infty}\frac{n_\Delta(-1)}{n_\Delta(+3)}=3, \tag{15}$$

and due to the absence of unbalanced triangles of the types shown in Figs. 1(b) and 1(d) in this phase, the frequencies of triads presented in Fig. 1 are $(\frac{1}{4},0,\frac{3}{4},0)$.

To verify our picture, we numerically investigated the frequencies f of different triangles in the network when we started the evolution of the system from various initial conditions. These frequencies are presented in Fig. 6. When we start from $\{x_{ij}\}=0$ then for low temperatures [see Fig. 6(b)] the observed fraction $f(+3)$ of triangles of the type shown in Fig. 1(a) is three times lower than the fraction $f(-1)$ of triangles of the type shown in Fig. 1(c) as predicted by Eq. (15). The fractions of triangles presented in Fig. 1 change abruptly from $(\frac{1}{4},0,\frac{3}{4},0)$ to $(\frac{1}{8},\frac{3}{8},\frac{3}{8},\frac{1}{8})$ at a critical temperature T^d (see Figs. 6(c) and 6(d); cf. also Fig. 10 in Ref. [30]). The latter distribution of various types of triads is also kept in high temperature when we start the temporal system evolution from the paradise state [see Fig. 6(h)] and it corresponds to the probabilities of three, two, one, and zero successes in three Bernoulli trials when the probability of success is equal to $\frac{1}{2}$.

The above results mean that the polarized state with two hostile cliques exists only at low temperatures and disappears abruptly at $T=T^d$. Let us stress that cliques emerge from random initial conditions in our numerical simulations. A similar discontinuous transition was observed in CTD dynamics in Refs. [13,14] when the initial density of positive links and the average $\{x_{ij}(t=0)\}$ were changed in a continuous way. The analytical approach developed in Ref. [14] shows that the critical value of $\{x_{ij}(t=0)\}$ for this transition should be equal to zero. When this average is smaller, then the system reaches an equilibrium in the two cliques state, and above this threshold the system evolves towards the paradise. Numerical simulations for the CTD model show a slightly higher value of this threshold [14]. This discontinuous transition can be understood using our mean-field approach developed in Sec. III as a result of the multistability of the system presented in Fig. 2. In fact, the unstable solution x^u of Eq. (7) in Fig. 2 divides the set of initial conditions into two basins of attraction [37]. When $\{x_{ij}(t=0)\}$ is below x^u the system evolves towards the solution $x^0=0$ and for an initial condition above this curve the system evolves towards x^s. Since our model at $T=0$ corresponds to the CTD model and the separatrix x^u in Fig. 2 approaches zero at this limit, the discontinuous disappearance of the polarized two-clique phase observed in Ref. [14] can be explained by crossing the separatrix curve. This critical behavior is equivalent to the bifurcation diagram corresponding to the so-called *fold catastrophe* [38] and similar discontinuous phase transitions resulting from system bistability have been observed in many systems, including pairs of weakly interacting Ising networks [39], a majority model of a network with communities [40], and an activity-driven temporal bilayer echo-chamber system [41].

VI. CONCLUSIONS

In this paper we present a simple analytical approach which describes the first-order phase transition observed in thermalized Heider's balance systems on the complete graphs. The proposed mean-field approximation predicts that the critical temperature of the system is equal to $T^c=(N-2)/a^c$, where N is the number of graph nodes, and $a^c\approx 1.716\,49$. This temperature corresponds to a discontinuous transition from a critical state $\{x_{ij}\}=x^c\approx 0.796\,388$ (that is close to the paradise state) to an unbalanced state with the same number of positive and negative links $\{x_{ij}\}=x^0=0$. At the critical point (T^c,x^c) the upper stable solution $x^s(T)$ for a fixed point of Eq. (8)—corresponding to a state of paradise dressed with thermal fluctuations at a given temperature T—coincides with the unstable branch $x^u(T)$ that is a separatrix, i.e., a boundary

between initial conditions leading to the "nearly paradise" solution $x^s(T)$ or to the disordered solution x^0. For $T > T^c$ the solution $x^s(T)$ does not exist anymore and the system is always evolving towards x^0. This bifurcation scenario corresponds to the well-known fold catastrophe [see Figs. 4(a) and 4(c)].

At the critical point, the system energy density changes from $\overline{E} = -(x^c)^3$ to the value $\overline{E} = 0$ [see Figs. 4(b) and 4(d)]. The results of computer simulations agree within the estimated uncertainties with our analytical calculations providing that initial conditions are close to the paradise state $\{x_{ij}(t = 0)\} = 1$.

On the other hand, at any temperature T when starting from a randomly selected state $x_{ij} = \pm 1$—when $\{x_{ij}(t = 0)\} = 0$—we reach in the simulation only the solution $\{x_{ij}\} = 0$ corresponding to the stable fixed point x^0. The solution can correspond to various patters of polarizations x_{ij}. At low temperatures, there is a phase consisting of two cliques of similar sizes that possess only positive internal links, but all interclique links are negative. It follows that there are triangles consisting of all positive links or triads of one positive link and two negative links. This means that all the triads are balanced and that the energy density of the system is the same as that in the paradise state, $\overline{E} = -1$; that is, degeneration of the ground is observed. The signatures of such system division—i.e., the ratio $f(-1)/f(+3) = 75\%/25\%$—for low-temperature noise level limit were also observed for diluted and densified triangulations (see Fig. 10 in Ref. [30]) and classical (Erdős-Rényi) random graphs (see Fig. 7 in Ref. [31]).

At a certain temperature $T = T^d$ [for $\{x_{ij}(t = 0)\} = 0$] or $T = T^c > T^d$ [for $\{x_{ij}(t = 0)\} = 1$], another discontinuous phase transition occurs when the number of unbalanced triads [Figs. 1(b) and 1(d)] abruptly increases, and as a result the system energy density becomes zero [see dashed arrows connecting open symbols in Figs. 4(b) and 4(d)]. This transition is not seen at a value of $\{x_{ij}\}$, which is 0 below and 0 above T^d when $\{x_{ij}(t = 0)\} = 0$.

We stress that for $T^d < T < T^c$—depending on the initial conditions—the system energy is either close to the ground state or equal to 0 [see Figs. 4(b) and 4(d)]. Assuming initial conditions $\{x_{ij}(t = 0)\} = 1$ and heating the system from $T = 0$ to $T = T^c$ decreases the mean value of $\{x_{ij}\}$ along the curve $x^u(T)$ to the critical point x^c, but above T^c we can only reach $\{x_{ij}\} = 0$. The observed transition is irreversible, and cooling the system from $T > T^c$ towards $T \to 0^+$ will never reproduce positive values of $\{x_{ij}\} > 0$, and in this sense the hysteresislike loop can be observed in the system; thus, (T_c, x_c) is a tipping point of our phase diagram [38].

The critical temperature $T^c = 26.2$ for the complete graph with $N = 50$ nodes estimated in Ref. [25] is roughly in agreement with our estimate of T^c for $N = 50$. Finding analytically the value of a lower critical temperature T^d—where, starting with $\{x_{ij}(t = 0)\} = 0$, a special mixture of only balanced triangles of the types shown in Figs. 1(a) and 1(c) disappears—is beyond predictions of our approach, and it remains a challenge.

The most interesting result of these investigations for real social networks could be the observation that, if such networks are driven by structural balance dynamics, then the balanced bipolar state seems to be the only possible state when the initial configuration of social links is completely random and the strengths of social interactions increase over time. To prohibit the emergence of such a polarized state, one could consider introducing additional attributes of interacting agents [42].

Finally, we note that Heider's theory may be applied for studies of international relations. In Ref. [13], the authors presented the evolution of relations between countries as a prelude to World War I. Further analysis of historical data is in progress [43].

ACKNOWLEDGMENTS

We are grateful to Krzysztof Kułakowski, Krzysztof Suchecki, and Piotr Górski for fruitful discussion and critical reading of the manuscript. This research has received funding as the RENOIR Project from the European Union's Horizon 2020 research and innovation program under Marie Skłodowska–Curie Grant No. 691152 and from the Ministry of Science and Higher Education (Poland) under Grants No. W34/H2020/2016 and No. 329025/PnH/2016. J.A.H. was also supported by POB Research Centre Cybersecurity and Data Science of Warsaw University of Technology within the Excellence Initiative Program—Research University (IDUB) and by the Polish National Science Center under Alphorn Grant No. 2019/01/Y/ST2/00058.

APPENDIX: LINK-LINK CORRELATIONS

In the mean-field approximation developed in Ref. [25] the mean polarization of links is given as

$$p = \langle x_{ij} \rangle = \tanh(\beta(N-2)q), \tag{A1}$$

where $\beta = 1/T$ and $q = \langle x_{ik} x_{kj} \rangle$ is a correlation between links and

$$q = \frac{e^{2(N-3)\beta q} - 2e^{-2\beta p} + e^{-2(N-3)\beta q}}{e^{2(N-3)\beta q} + 2e^{-2\beta p} + e^{-2(N-3)\beta q}}. \tag{A2}$$

Then the values p and q were found numerically as functions of the temperature T from Eqs. (A1) and (A2) in Ref. [25]. In our case, we receive the mean-field solution from the fixed point of Eq. (7), that is, $\langle x_{ij} \rangle = \tanh(\beta(N-2)\langle x_{ij} \rangle^2)$. However, one can easily find that when $N \gg 1$ then

$$p^2 \approx q. \tag{A3}$$

In fact, in the thermodynamic limit, the difference between q and p^2 is due to the second term of the nominator of Eq. (A2), which is $z = 2\exp(-2\beta p)$ and should be equal to 2 if Eq. (A3) is valid. However, one can write $z = 2\exp[-(2p/T^c)(T^c/T)]$, where T^c is a critical temperature, and since every link is influenced by many triangles in the Hamiltonian (2), thus $T^c \gg 1$. Thus, when the system is in the critical region and $T \approx T^c$, and since $p/T^c \ll 1$, thus $z \approx 2$ and $p \approx q^2$. The approximation also works well in the low-temperature region $T \to 0^+$ since in such a case the terms z and 2 are much smaller than the first term in the nominator of Eq. (A2) and $q = 1$.

[1] F. Heider, Attitudes and cognitive organization, J. Psychol. **21**, 107 (1946).

[2] F. Harary, On the notion of balance of a signed graph, Mich. Math. J. **2**, 143 (1953).

[3] D. Cartwright and F. Harary, Structural balance: A generalization of Heider's theory, Psychol. Rev. **63**, 277 (1956).

[4] F. Harary, On the measurement of structural balance, Behav. Sci. **4**, 316 (1959).

[5] J. A. Davis, Clustering and structural balance in graphs, Hum. Relat. **20**, 181 (1967).

[6] F. Harary, R. Z. Norman, and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs*, 3rd ed. (Wiley & Sons, New York, 1965).

[7] P. Doreian and A. Mrvar, A partitioning approach to structural balance, Soc. Networks **18**, 149 (1996).

[8] J. A. Rambaran, J. K. Dijkstra, A. Munniksma, and A. H. N. Cillessen, The development of adolescents' friendships and antipathies: A longitudinal multivariate network test of balance theory, Soc. Networks **43**, 162 (2015).

[9] C. M. Rawlings and N. E. Friedkin, The structural balance theory of sentiment networks: Elaboration and test, Am. J. Sociol. **123**, 510 (2017).

[10] Y.-S. Chiang and L. Tao, Structural balance across the strait: A behavioral experiment on the transitions of positive and negative intergroup relationships in mainland China and Taiwan, Soc. Networks **56**, 1 (2019).

[11] A. Toroslu and E. Jaspers, Avoidance in action: Negative tie closure in balanced triads among pupils over time, Soc. Networks **70**, 353 (2022).

[12] A. Bramson, K. Hoefman, K. Schoors, and J. Ryckebusch, Diplomatic relations in a virtual world, Political Anal. **30**, 214 (2022).

[13] T. Antal, P. L. Krapivsky, and S. Redner, Dynamics of social balance on networks, Phys. Rev. E **72**, 036121 (2005).

[14] T. Antal, P. L. Krapivsky, and S. Redner, Social balance on networks: The dynamics of friendship and enmity, Phys. D (Amsterdam, Neth.) **224**, 130 (2006).

[15] K. Kułakowski, P. Gawroński, and P. Gronek, The Heider balance: A continuous approach, Int. J. Mod. Phys. C **16**, 707 (2005).

[16] S. A. Marvel, J. Kleinberg, R. D. Kleinberg, and S. H. Strogatz, Continuous-time model of structural balance, Proc. Natl. Acad. Sci. USA **108**, 1771 (2011).

[17] P. Gawroński, M. J. Krawczyk, and K. Kułakowski, Emerging communities in networks—A flow of ties, Acta Phys. Pol., B **46**, 911 (2015).

[18] P. J. Górski, K. Bochenina, J. A. Hołyst, and R. M. D'Souza, Homophily Based on Few Attributes Can Impede Structural Balance, Phys. Rev. Lett. **125**, 078302 (2020).

[19] K. Malarz, M. Wołoszyn, and K. Kułakowski, Towards the Heider balance with a cellular automaton, Phys. D (Amsterdam, Neth.) **411**, 132556 (2020).

[20] K. Malarz and K. Kułakowski, Heider balance of a chain of actors as dependent on the interaction range and a thermal noise, Phys. A (Amsterdam, Neth.) **567**, 125640 (2021).

[21] M. J. Krawczyk, K. Kułakowski, and Z. Burda, Towards the Heider balance: Cellular automaton with a global neighborhood, Phys. Rev. E **104**, 024307 (2021).

[22] Z. Burda, M. J. Krawczyk, and K. Kułakowski, Perfect cycles in the synchronous Heider dynamics in complete network, Phys. Rev. E **105**, 054312 (2022).

[23] M. J. Krawczyk and K. Kułakowski, Structural balance in one time step, Phys. A (Amsterdam, Neth.) **606**, 128153 (2022).

[24] E. Estrada and M. Benzi, Walk-based measure of balance in signed networks: Detecting lack of balance in social networks, Phys. Rev. E **90**, 042802 (2014); R. Nishi and N. Masuda, Dynamics of social balance under temporal interaction, Europhys. Lett. **107**, 48003 (2014); V. B. Kovchegov, Modeling of human society as a locally interacting product-potential networks of automaton, arXiv:1605.02377; P. J. Górski, K. Kułakowski, P. Gawroński, and J. A. Hołyst, Destructive influence of interlayer coupling on Heider balance in bilayer networks, Sci. Rep. **7**, 16047 (2017); M. J. Krawczyk, M. Wołoszyn, P. Gronek, K. Kułakowski, and J. Mucha, The Heider balance and the looking-glass self: Modelling dynamics of social relations, *ibid.* **9**, 11202 (2019); S. Sheykhali, A. H. Darooneh, and G. R. Jafari, Partial balance in social networks with stubborn links, Phys. A (Amsterdam, Neth.) **548**, 123882 (2020); T. M. Pham, A. C. Alexander, J. Korbel, R. Hanel, and S. Thurner, Balance and fragmentation in societies with homophily and social balance, Sci. Rep. **11**, 17188 (2021); H. Deng, M. Qi, M. Li, and B. Ge, Limited cognitive adjustments in signed networks, Phys. A (Amsterdam, Neth.) **584**, 126325 (2021); S. Arabzadeh, M. Sherafati, F. Atyabi, G.R. Jafari, and K. Kułakowski, Lifetime of links influences the evolution towards structural balance, *ibid.* **567**, 125689 (2021); M. H. Hakimi Siboni, A. Kargaran, and G. R. Jafari, Hybrid balance theory: Heider balance under higher-order interactions, Phys. Rev. E **105**, 054105 (2022); M. J. Krawczyk, M. del Castillo-Mussot, E. Hernandez-Ramirez, G. G. Naumis, and K. Kułakowski, Heider balance, asymmetric ties, and gender segregation, Phys. A (Amsterdam, Neth.) **439**, 66 (2015); K. Kułakowski, M. Stojkow, and D. Żuchowska-Skiba, Heider balance, prejudices and size effect, J. Math. Sociol. **44**, 129 (2020); M. J. Krawczyk and K. Kułakowski, Structural balance of opinions, Entropy **23**, 1418 (2021).

[25] F. Rabbani, A. H. Shirazi, and G. R. Jafari, Mean-field solution of structural balance dynamics in nonzero temperature, Phys. Rev. E **99**, 062302 (2019).

[26] R. Shojaei, P. Manshour, and A. Montakhab, Phase transition in a network model of social balance with Glauber dynamics, Phys. Rev. E **100**, 022303 (2019).

[27] K. Malarz and K. Kułakowski, Comment on "Phase transition in a network model of social balance with Glauber dynamics," Phys. Rev. E **103**, 066301 (2021).

[28] P. Manshour and A. Montakhab, Reply to "Comment on 'Phase transition in a network model of social balance with Glauber dynamics," Phys. Rev. E **103**, 066302 (2021).

[29] K. Malarz and M. Wołoszyn, Expulsion from structurally balanced paradise, Chaos **30**, 121103 (2020).

[30] M. Wołoszyn and K. Malarz, Thermal properties of structurally balanced systems on diluted and densified triangulations, Phys. Rev. E **105**, 024301 (2022).

[31] K. Malarz and M. Wołoszyn, Thermal properties of structurally balanced systems on classical random graphs, arXiv:2206.14226.

[32] K. Kułakowski, A note on temperature without energy—A social example, arXiv:0807.0711.

[33] K. Binder, Applications of Monte Carlo methods to statistical physics, Rep. Prog. Phys. **60**, 487 (1997).

[34] T. T. Wu, Theory of Toeplitz determinants and the spin correlations of the two-dimensional Ising model. I, Phys. Rev. **149**, 380 (1966).

[35] B. N. Taylor and C. E. Kuyatt, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297 (National Institute of Standards and Technology, Gaithersburg, MD, 1994).

[36] M. J. Krawczyk, S. Kaluzny, and K. Kułakowski, A small chance of paradise—Equivalence of balanced states, Europhys. Lett. **118**, 58005 (2017).

[37] E. Ott, *Chaos in Dynamical Systems*, 2nd ed. (Cambridge University, Cambridge, England, 2002); K. Kacperski and J. A. Hołyst, Phase transitions and hysteresis in a cellular automata-based model of opinion formation, J. Stat. Phys. **84**, 169 (1996).

[38] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara, Early-warning signals for critical transitions, Nature (London) **461**, 53 (2009).

[39] K. Suchecki and J. A. Hołyst, Bistable-monostable transition in the Ising model on two connected complex networks, Phys. Rev. E **80**, 031110 (2009); Ising model on two connected Barabasi-Albert networks, **74**, 011122 (2006).

[40] R. Lambiotte, M. Ausloos, and J. A. Hołyst, Majority model on a network with communities, Phys. Rev. E **75**, 030101(R) (2007).

[41] Ł. G. Gajewski, J. Sienkiewicz, and J. A. Hołyst, Transitions between polarization and radicalization in a temporal bilayer echo-chamber model, Phys. Rev. E **105**, 024125 (2022).

[42] P. J. Górski, C. Atkisson, and J. A. Hołyst, A general model for how attributes can reduce polarization in social groups (unpublished).

[43] E. Estrada, F. Diaz-Diaz, and P. Bartesaghi, Local balance index in signed networks (unpublished).