Scilight

Journal of Social Physics
https://www.sciltp.com/journals/jsp

Article

Dynamics of Social Balance with Ternary Interpersonal Relationships

Hirotaka Goto [1,2,*], Masashi Shiraishi [3,4], Hiraku Nishimori [4] and Joshua B. Plotkin [2,5]

[1] Graduate School of Advanced Mathematical Sciences, Meiji University, 4-21-1 Nakano, Tokyo 164-8525, Japan
[2] Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA
[3] Graduate School of Information Sciences, Hiroshima City University, 3-4-1 Ozukahigashi, Asaminami-ku, Hiroshima 731-3194, Japan
[4] Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, 4-21-1 Nakano, Tokyo 164-8525, Japan
[5] Center for Mathematical Biology, University of Pennsylvania, Philadelphia, PA 19104, USA
* Correspondence: hirotakagoto@meiji.ac.jp

How To Cite: Goto, H.; Shiraishi, M.; Nishimori, H.; et al. Dynamics of Social Balance with Ternary Interpersonal Relationships. *Journal of Social Physics* **2025**, *1*(1), 1.

Received: 18 October 2025
Revised: 10 December 2025
Accepted: 12 December 2025
Published: 18 December 2025

Abstract: Interpersonal relationships are building blocks of human social groups and resulting social behavior. The dynamics of social groups can be critical in understanding the origins of polarization, cooperation, and conflict. *Social balance* is a classical notion that defines which configurations of pairwise relationships among three individuals are favored. Previous studies have found that the principle of social balance, when applied to all triads within a population, leads to two mutually antagonistic groups that are internally friendly or one giant friendly group. However, this body of theory assumes interpersonal relationships are binary—either friendly or unfriendly. *Neutrality* adds another layer of complexity. We develop and analyze two models of social balance with interpersonal relationships that are positive, negative, or neutral. We find that the population almost never reaches a fully polarized or paradise state: even though the overall frequencies of triad configurations reach an equilibrium, individuals continue to update their interpersonal relationships over time. We also identify the equilibrium triad frequencies, based on a mean-field analysis. Subsequently, we derive an effective mapping between these two alternative models, which clarifies a nontrivial connection between distinct social processes. We also find a phase transition when some parameters are small, which characterizes the emergence of a giant component connected by active (i.e., friendly or unfriendly) links; and we discuss how the critical regime is relevant in a growing society. Our analysis helps fill a theoretical gap in the literature, and reveals qualitatively new dynamics that arise from a theory of triadic social interactions.

Keywords: interpersonal relationships; balance theory; signed networks; complex networks; triadic interactions

1. Introduction

Friends often know each other. This is certainly not a coincidence: we tend to meet new people through mutual friends. Having common enemies also tends to make our bonds stronger. When two individuals get along, it is also perhaps because one of them has noticed the other being on good terms with some other person, whether a stranger or a friend of theirs; after all, reputations are critical in building new social relationships in most societies. In short, our friendship circles grow and evolve not necessarily based on pairwise interactions but often through third parties; therefore, the connections become correlated in one way or another.

Empirical friendship networks have been studied for decades [1]. Over the past few decades, the interplay between networks and social dynamics has become a central theme in the study of social physics, as the amount of available large-scale data has increased rapidly. For example, opinion dynamics on networks is a cornerstone of this

Copyright: © 2025 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

research area [2–6]. In parallel with these directions, research has also shown how networks of interpersonal relations develop and organize themselves, especially in online settings, which has focused scientific attention on the dynamics *of* underlying (signed) social networks themselves—not just how information and opinions change *on* them. Empirical examples include the study of the famous karate club network [7] and many other social and online networks [8–14], also extending to international relations [15,16], workplace environments [17], and even fairy tales [18]. Signed networks also seem to play a role in ecological contexts, such as inter-species relationships in plants [19] and social network structures in mammals [20], as well as in pet-related services [21]. The most canonical theory overarching these studies is the so-called *balance theory*.

Balance theory, which dates back to the original work by psychologist Heider [22], posits how friendships or enmities between three individuals should be structured. This basic theory has led to extensive theoretical and empirical work on such triadic interactions. The theory is based on three fundamental premises: relationships are *mutual*; in binary form (either *friendly* or *unfriendly*); and a group of three individuals formed by three mutual relationships (i.e., a triad) is either *balanced* or *imbalanced*. The theory posits that a triad with an odd number of friendships (one or three) is balanced, whereas one with an even number (zero or two) is imbalanced. The balanced triads are assumed to be favored over the imbalanced ones, perhaps through some dynamic process that updates one of the pairwise relationships from either friendly to unfriendly, or vice versa. In short, the theory of social balance succinctly captures the intuition behind sayings like "a friend of a friend is also a friend" and "an enemy of a friend is also an enemy".

When the simple principle of social balance is extended to a complete graph, it predicts that a signed graph is balanced either when all edges are friendly (paradise) or when it partitions the nodes into two subsets that are internally friendly but mutually unfriendly (polarized state) [23,24]. This classical theory is based on a *static* view, i.e., only considering whether or not, and how many of the triads, are balanced in a given network. More recently, dynamical models of social balance have been introduced, where the number of imbalanced triads decreases as link signs flip during every update event, until no imbalanced triads remain. Antal et al. explored two such discrete dynamical models in their seminal work [25]. Particularly relevant to our study is the model of "local triad dynamics", where a randomly chosen triad is updated based on its sign configuration. In a finite population, this process eventually achieves complete balance, regardless of initial conditions, producing either a paradise state (everyone is friends with everyone else) or a polarized state (two subgroups are each internally friendly but unfriendly to each other), although the time required to reach these equilibria can grow rapidly with population size. Notably, a mean-field approximation alone cannot predict the long-term polarized outcome in a finite population, because full polarization arises from symmetry breaking and it violates the mean-field assumption. Therefore, researchers have applied the theory of random walks to characterize the long-term dynamical behavior, including the absorption time for reaching balance.

Those previous models demonstrate dynamical pathways to complete balance, which is a satisfying result as it provides a clear basis for understanding how balance can be achieved from random initial conditions. Nevertheless, discrete models lack a mechanism that can describe gradual changes in the strengths of interpersonal relationships. To this end, continuous models (both in time and state) have been proposed, which can account for incremental changes in relationship strengths driven by the tendency to achieve higher overall balance [26–28]. Even in the continuous setting, though, a fully balanced state is still the primary outcome, regardless of the model details. In other words, even when links can assume any real value, these models are constrained in such a way that the interactions end up in complete balance, i.e., they are still within the realm of classical social balance theory.

There has been an attempt to extend the classical framework of balance theory to ternary interpersonal relationships—that is, relationships that can be either positive, negative, or inactive [29]. Extending balance theory in this direction deserves more attention because triads with all three links active (either positive or negative) are much less frequent in empirical data. For example, in political networks, such fully active triads comprise only a few percent of all triads [29]. Moreover, from a theoretical perspective, adding neutral links to a signed networks provides an analytical advantage by making the network a complete graph [29].

A recent study has also found that neutrality plays a pivotal role in international relations and argues that the traditional balance theory should be reformulated to accommodate this effect [30]. The "agentic zero" is a neutral relationship between two parties, also referred to as the "Switzerland option". The absence of ties may arise as a strategic choice, or by a convention that discourages tension, or simply because forming ties is costly. While the role of null ties is commonly addressed in studies using blockmodeling [31], they are, in the study of social balance, often dismissed as "missing data" or sometimes treated as negative ties by assuming that a lack of a positive sentiment implies a negative one [30]. It would seem preferable to extend the theory of social balance to incorporate the possibility of neutral ties, including in dynamical theories of link formation and dissolution.

When pairwise links are ternary (positive, negative, or neutral), we must consider ten possible configurations of triads (see Figure 1) and develop a new, revised concept of balance suitable under this extended framework. One model proposed in a previous study [29], which combines both theoretical and empirical approaches, is based on a prescribed Hamiltonian, consisting of three-edge, two-edge, and one-edge interaction terms as well as a "chemical potential" term. This means that it assumes energy levels *a priori* to determine the "stability" of all triads. However, such notions of "social stability" would benefit from some mechanistic explanation, especially one that accounts for how they arise from individual-level interactions and decision-making in a bottom-up manner. Therefore, in this paper we aim to develop a theory of social balance with three-state interpersonal relationships, using a framework based on microscopic, individual-level stochastic processes. This can be seen as a direct extension of the work by Antal et al. [25] to accommodate neutral, as well as positive and negative, interpersonal relationships.

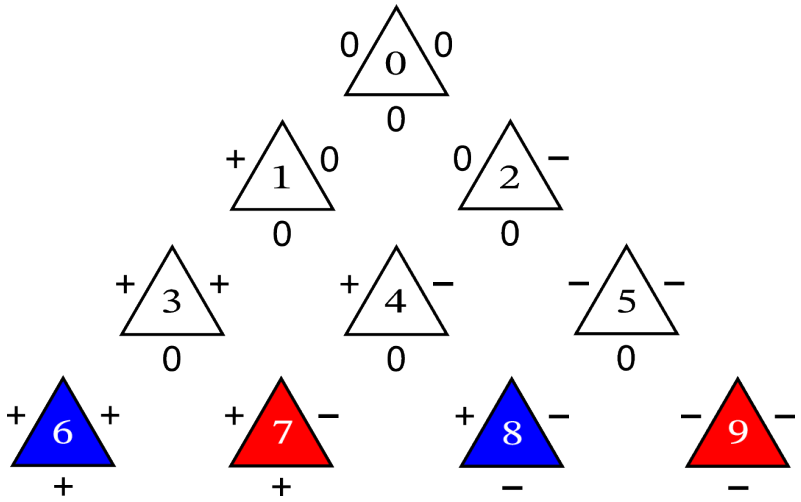

Figure 1. Ten types of triads in our extended social balance theory, where a third type of interpersonal relationships is introduced in addition to the traditional friendly and unfriendly—namely, the *neutral* (or inactive) relationship.

Here, we develop two models of social balance that accommodate neutral relations. We study the equilibrium states of these models using a mean-field analysis and associated approximations. As a result, we discover that these alternative models can be mapped to each other, under certain parameter conditions. This is noteworthy especially because it uncovers a nontrivial relationship between distinct social processes. Importantly, we find that polarized communities generally do not arise in either of these models. This implies that polarization in the classical case might be in part attributed to the binary representation of interpersonal relationships. As opposed to rapid absorption into balance, we observe transient dynamics at the microscopic scale (constantly "rewiring" links), while mesoscopic quantities fluctuate around equilibrium values. Quite different from the classical model, we also identify a phase transition that characterizes the emergence of a giant component connected by active links at equilibrium. We then highlight the tradeoff between these two regimes, one in which a giant component emerges and the other in which it does not. On the one hand, forming a giant component, roughly of the size of the population, requires many instances of activating an inactive link, which in real life should be costly—especially so when the population size is significantly large. At the other end of the spectrum, a group of isolated individuals would not function properly as a society. Therefore, the in-between scenario, i.e., the critical point of the phase transition, makes the best of both, offering a robust advantage for societies to ensure both feasibility and marginal network connectivity, even when the population size is large. Our results provide a basis for a qualitatively new, and empirically motivated, extension to social balance, suggesting new directions for research at the interface of balance theory, complex networks, and social dynamics.

2. Models

2.1. Focal-Link Model

We consider an undirected network of N individuals fully connected via $L = \binom{N}{2}$ pairwise relationships. We represent individuals and pairwise relationships by nodes and links, respectively. Each link is assigned a discrete value at each timestep t, denoted $x_{ij}(t) \in \{-1, 0, 1\}$ (note that $x_{ij}(t) = x_{ji}(t)$ for all $i, j \in \{1, \ldots, N\}$). In this ternary setting, triads Δ_6 and Δ_8 are the only balanced configurations, as in the classical setting.

During each update step, we randomly pick a focal link $\{i, j\}$ and an individual k so that all three individuals are distinct. Then, we update, or "reassess," x_{ij} based on the current signs of x_{ik} and x_{kj}. If both x_{ik} and x_{kj} are

zero, we set either $x_{ij} = 1$ or $x_{ij} = -1$ with equal probability ϵ; we set it to zero with the remaining probability $1 - 2\epsilon$. If exactly one of x_{ik} and x_{kj} is positive (negative), we set $x_{ij} = 1$ ($x_{ij} = -1$) with probability δ; we set it to $x_{ij} = 0$ with the remaining probability $1 - \delta$. If both x_{ik} and x_{kj} are nonzero, we update x_{ij} with probability η in such a way that the updated triad is balanced; we set it to zero with the remaining probability $1 - \eta$. We randomly and independently pick a focal link and associated triad, and then update (or reassess) the link, L times during a unit time. This ensures that on average every link is updated once per unit time. Figure 2 illustrates the definition of the above mentioned update events, each of which we refer to as an ϵ-, δ-, or η-event. The parameters are restricted to the ranges $\epsilon \in (0, 1/2]$ and $\delta, \eta \in [0, 1]$. Note that we have omitted the case $\epsilon = 0$ to avoid ill-posedness. Because the update process focuses on link reassessment, we refer to this model as the *focal-link* model.

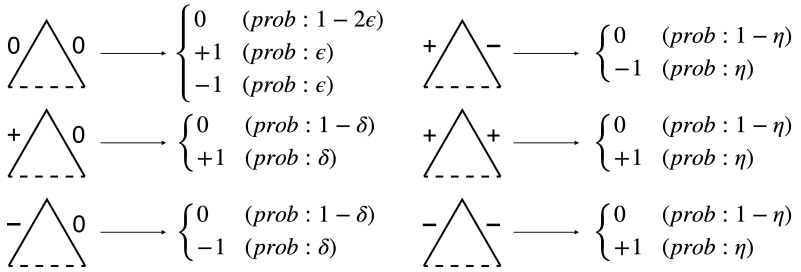

Figure 2. Update rules for the focal-link model. The solid and dashed lines represent a focal and its neighboring links, respectively. These update events are called either ϵ-, δ-, or η-events, depending on the number of active neighboring links within the chosen triad.

Let us briefly interpret the parameters of the focal-link model (see Table 1). In what follows, we call a link *inactive* if it is zero; otherwise, *active*. Similarly, an *active (inactive) triad* has all three links active (inactive), whereas a *semi-active triad* has one or two links active. Furthermore, given a chosen triad with a focal link, we refer to the other two links forming the triad together as the *neighboring* links. The parameter ϵ represents the likelihood of activating a focal link after reassessing it—regardless of its current status—when both neighboring links are inactive. It is therefore the probability that a focal link gets activated when it is reassessed based on the current statuses of the two inactive neighboring links. The parameter δ indicates the likelihood that a focal link becomes active, with the same sign as that of the only active neighboring link. We refer to this process as (signed) *preferential link reassessment*. In other words, δ is the probability that the only active neighboring link activates the focal link, i.e., how likely it "attracts" another active link. The parameter η represents the likelihood that two active neighboring links—regardless of their signs—lead to a balanced triad by adjusting the sign of the focal link. That is, with probability η, a *balanced* state is favored over a *relaxed* one in which the focal link is inactive.

The focal-link model consists of only three parameters and essentially three types of events (namely, ϵ-, δ-, and η-events), and therefore is developed to be as concise as possible. Nevertheless, in exchange for simplicity, it suffers an apparent limitation: every update is based purely on reassessment of a focal link, without accounting for the current sign of the focal link. This model can produce somewhat unrealistic behavior when the results are interpreted in a social context. In particular, the focal-link model allows for a positive link to turn negative and vice versa in a *single* update event, even when the triad is not fully active (i.e., when the triad includes at least one inactive link). We call such events *irrational updates* (see Figure 3 for an example; see also the horizontal transitions in the middle two rows in Figure S1). In a real-life social scenario, by contrast, it seems more reasonable for the inactive link in such a triad to become active and produce a balanced triad. To address the deficiency of irrational updates in the focal-link model, we next develop an alternative model, called the focal-triad model.

Table 1. Description of the parameters.

Model	Parameter	Description
Focal-link	ϵ	Random reassessment
	δ	Preferential reassessment (signed)
	η	Tendency toward balance
Focal-triad	$\hat{\epsilon}$	Random activation
	κ_1	Social contagion
	κ_2	Social isolation
	$\hat{\eta}$	Tendency toward balance

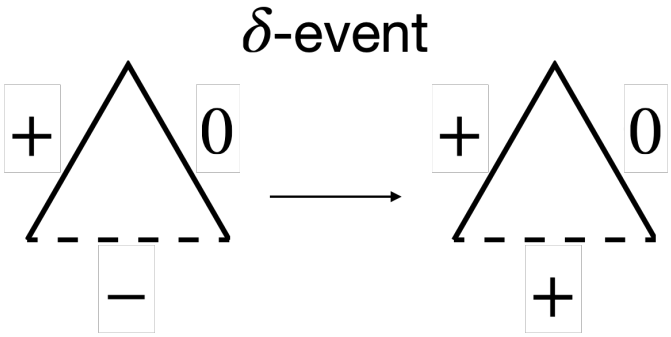

Figure 3. An example of an *irrational update* via a δ-event, in which Δ_4 transitions to Δ_3 under the focal-link model. The dashed and solid lines represent the focal-link and its neighboring links, respectively. The focal link flips sign as a result of a preferential reassessment event. We call this type of event irrational because it is a complete reversal of link's sign that does not produce a balanced triad. Such irrational updates can also occur between triads Δ_1 and Δ_2, between Δ_3 and Δ_4, and between Δ_4 and Δ_5 in the focal-link model (see Supplemenatry Materials Figure S1). The focal-triad model removes the possibility of irrational updates to a focal triad (see Figure 4).

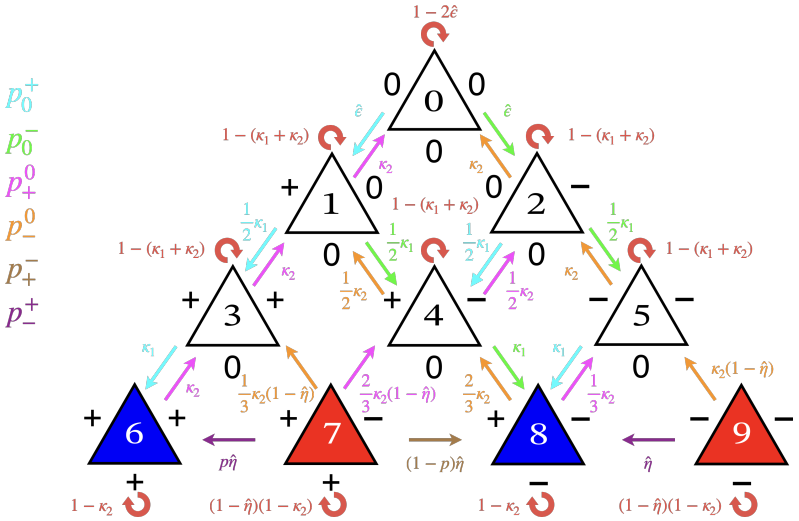

Figure 4. Update rules for the focal-triad model. In contrast to the focal-link model (Figure 2), the focal-triad model directly specifies the probabilities of transitioning between different types of triads using four parameters: $\hat{\epsilon}$, κ_1, κ_2, and $\hat{\eta}$. Note that $\kappa_1 + \kappa_2 \leq 1$ is required. The parameter p corresponds to the propensity for friendly links in Ref. [25]; in this paper, we fix $p = 1/3$. Irrational updates do not occur within any focal triad in the focal-triad model, whereas they can occur in the focal-link model (cf. Supplemenatry Materials Figure S1).

2.2. Focal-Triad Model

Now we develop our second model, where updates are based on *triads* as opposed to *links*—i.e., it is defined directly through the transition probabilities among the ten triad types. Figure 4 illustrates the definition of our second model, where four parameters, $\hat{\epsilon} \in (0, 1/2]$, $\kappa_1, \kappa_2 \in [0, 1]$, and $\hat{\eta} \in [0, 1]$, are introduced. In each update event, we pick a focal triad at random and change one of its link signs based on the transition probabilities specified in the diagram. We iterate this event L times, which constitutes a unit time. We call this model the *focal-triad* model.

In line with our expectation, the focal-triad model, by design, successfully reduces the frequency of such irrational updates (see Section 3.2.3 and S5 in Supplemenatry Materials for details), although it comes at the cost of an increased number of model parameters. Finally, we note that horizontal transitions at the bottom row should not be dismissed because all those triads are fully active, hence the classical balance theory holds (see Figure 4).

Whereas the focal-link model has the advantage of fewer parameters, the focal-triad model allows for a more comprehensible interpretation of its parameters. In particular, $\hat{\epsilon}$ is the probability of activating a fully inactive triad; κ_1 is the probability of further activating a triad that is only partially active (semi-active), a process we refer to as *social contagion*; κ_2 is the probability of inactivating a fully or semi-active triad, a process we call *social isolation*; and $\hat{\eta}$ is the probability that an imbalanced triad favors transitioning to balance over inactivating one of its active links (see Table 1 for a complete list). In addition, this model formulation removes the possibility of an irrational update occurring in any focal triad.

3. Results

3.1. Equilibrium States

Figure 5 shows the time evolution of the frequencies of all types of triads, obtained by stochastic simulations, in both the focal-link and focal-triad models (dashed lines). Both models reach an equilibrium state after a sufficiently long period of time. Notably, the stationary values do not depend on the initial conditions; we suspect that there is only one stable equilibrium state that the system ends up in, irrespective of (non-pathological) initial conditions. To study this rigorously, we develop a mean-field theory by extending the method used in a previous study [25] (see Supplemenatry Materials for details). Our analytical treatment is as follows: we first derive closed-form rate equations from the stochastic process; under certain parameter conditions, we identify approximate equilibrium rescaled frequencies of all types of triads; and we confirm the results by solving the rate equations numerically and comparing them to stochastic simulations. We apply this method in both the focal-link and focal-triad models.

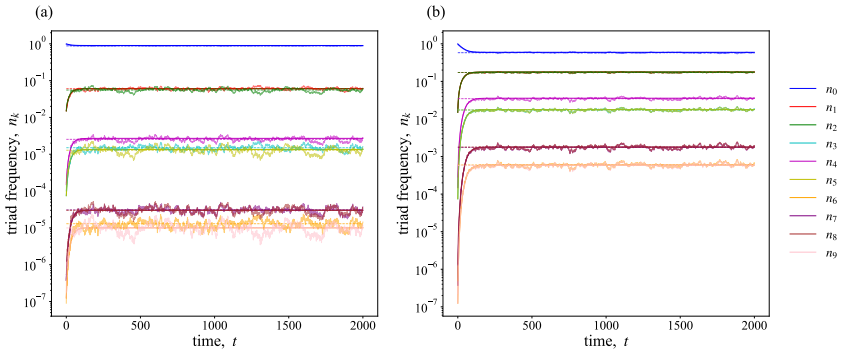

Figure 5. Time-series frequencies of all types of triads: stochastic process (solid, fluctuating), rate equations (solid, smooth), time-averaged data (dashed, straight). The results show that the symmetry ansatz, Equation (1), holds approximately at equilibrium in both models, given these parameter regimes. Parameters: (**a**) $\epsilon = 0.0005$, $\delta = 0.5$, $\eta = 0.5$ (focal-link model), (**b**) $\hat{\epsilon} = 0.0015$, $\kappa_1 = 0.5$, $\kappa_2 = 0.5$, $\hat{\eta} = 0.5$ (focal-triad model). Initially, most links are inactive, with a small equal fraction of positive and negative links. Note that in the stochastic models, a unit time consists of L update events. Since updating a link causes $N-2$ triads to be updated as well, each triad is updated three times on average per unit time ($L(N-2)/\binom{N}{3} = 3$). Therefore, we have rescaled time as $t \to 3t$ in the rate equations to maintain consistency with stochastic simulations.

Let Δ_k denote the type-k triad ($k = 0, \ldots, 9$), where the numbering is the same as in Figure 1. We represent its corresponding frequencies by n_k, which equilibrate eventually. In Supplemenatry Materials we derive rate equations that govern the time evolution of these ten variables. The mean-field predictions, obtained by numerically solving the rate equations, are in fact consistent with stochastic simulations, even in a relatively small population of size $N = 512$ (see Figure 5), confirming the validity of our analytical approach. This is important in particular because there is no guarantee that a mean-field approach using rate equations works as intended; indeed, it fails to predict the qualitative long-term behavior of a finite population (which includes a completely balanced outcome) in traditional models of social balance with binary relations [25]. In fact, the mean-field theory also fails in the classical scenarios (e.g., $\epsilon = 1/2$ and $\delta = \eta = 1$ in the focal-link model; $\hat{\epsilon} = 1/2$, $\kappa_1 = \hat{\eta} = 1$, and $\kappa_2 = 0$ in the focal-triad model), in agreement with the prior work. However, these classical scenarios appear to be the only cases in which the mean-field approximation fails in our models, i.e., measure-zero cases in the full parameter space.

To predict equilibrium triad frequencies, we should identify the stationary state of the rate equations. Without simplifying them further, however, it is almost certainly not possible to obtain analytical expressions because there are ten highly nonlinear coupled differential equations. Therefore, we develop an approximation method that can significantly reduce the complexity of the ODE system, given certain conditions on the parameters.

3.2. Mean-Field Analysis

We analyze the equilibrium frequencies using the rate equations derived in Supplemenatry Materials. Let $r_k = n_k/n_0$ denote the rescaled triad densities, where the index goes over $k = 1, \ldots, 9$. At equilibrium, when a majority of links are inactive—which we call the *activation-limited regime*—we assume sign symmetry in frequency, i.e., exchanging positive and negative links does not affect the frequencies of the triads that transform into each other under the exchange, that is,

$$n_1 = n_2, \quad n_3 = n_5, \quad n_6 = n_9, \quad n_7 = n_8. \tag{1}$$

This ansatz is based on the observation that transition probabilities between different types of triads are arranged in a symmetric way. These probabilities can be derived from the focal-link model, whereas the focal-triad model is directly specified by them (see Supplemenatry Materials Figure S1 for the focal-link model and Figure 4 for the focal-triad model). We expect Equation (1) to be approximately true in the activation-limited regime because (i) the transition probabilities between triad types are defined in a horizontally symmetric way (left-right symmetric) except in the bottom row of triads; and (ii) most links are inactive so the bottom row triads are infrequent. Consequently, we expect, in the activation-limited regime, that

$$n_0 \gg n_1, n_2 \gg n_3, n_4, n_5 \gg n_6, n_7, n_8, n_9, \tag{2}$$

i.e., there will be a hierarchy of equilibrium triad densities in which less frequent types of triads are most often generated from triads of more frequent types. We should also expect a hierarchy of relaxation times: n_0 reaches equilibrium faster than n_1 and n_2; n_1 and n_2 reach equilibrium faster than n_3, n_4, and n_5; and n_3, n_4, and n_5 reach equilibrium faster than n_6, n_7, n_8 and n_9.

Let $p_\sigma^{\sigma'}$ denote the probability that a link changes its sign from σ to σ' due to an update event. Since we assume sign symmetry at equilibrium, we also expect that the following relations hold:

$$p_0^+ = p_0^-, \quad p_+^0 = p_-^0, \quad p_+^- = p_-^+. \tag{3}$$

These originate from the assumption that triad frequencies are symmetric with respect to positive and negative signs. Provided that we have $p_\sigma^{\sigma'}$ for all $\sigma \neq \sigma'$ either derived from the model or by definition (see Supplemenatry Materials), Equations (1) and (3), taken together, imply that

$$2n_3 = n_4, \quad 3n_6 = n_7. \tag{4}$$

These equations hold in both the focal-link and focal-triad models, which suggests that the mean-field system can effectively be described by only four variables—namely, n_0, n_1, n_3, and n_6—in the activation-limited regime (see Supplemenatry Materials for details).

To summarize, our aim is to identify approximate solutions to the rate equations when a majority of links are inactive, assuming sign symmetry at equilibrium. Symmetry conditions in equilibrium triad frequencies and transition probabilities, as well as the hierarchy conditions in frequency and timescale, follow immediately from the assumption on sign symmetry at equilibrium in the activation-limited regime. In Supplemenatry Materials we show that this approach indeed yields reasonably good approximate solutions in a self-consistent way.

3.2.1. Focal-Link Model: Scaling Behavior

Next we analyze approximate equilibrium frequencies and their scaling behavior. First, the activation-limited regime requires $\epsilon \ll 1/2$. Then, we find that r_1 approximately satisfies the following quadratic equation (see Supplemenatry Materials for details):

$$4(1-2\epsilon)r_1^2 + 3(1-2\epsilon-2\delta)r_1 - 9\epsilon = 0, \tag{5}$$

whose positive solution is

$$r_1 = \frac{6\epsilon}{(1-2\epsilon-2\delta) + \sqrt{(1-2\epsilon-2\delta)^2 + 16(1-2\epsilon)\epsilon}}. \tag{6}$$

Notice that Equation (5) does not depend on η. This implies that in the activation-limited regime, η does not affect equilibrium frequencies of top row triads, namely, n_0, n_1, or n_2, up to leading order. In other words, when most links are zero, the classical balance theory contributes only a negligible amount to the equilibrium frequencies of triads of type Δ_0, Δ_1, and Δ_2—those in which more than half the links that form a triad are inactive.

We also expand Equation (5) in power series of ϵ, which leads to the following expression (see Supplemenatry Materials for a derivation):

$$r_1 = \begin{cases} \frac{3\epsilon}{1-2\delta} + O(\epsilon^2) & \delta < 1/2, \\ \frac{3(2\delta-1)}{4} + \left[\frac{3(2\delta-1)}{2} + \frac{1}{2} + \frac{3}{2\delta-1}\right]\epsilon + O(\epsilon^2) & \delta > 1/2. \end{cases} \tag{7}$$

Notably, Equation (7) implies that as $\epsilon \to 0$, r_1 tends to

$$r_1 \to \begin{cases} 0 & \delta \le 1/2, \\ \frac{3(2\delta-1)}{4} & \delta \ge 1/2, \end{cases} \tag{8}$$

implying a phase transition at $\delta = 1/2$ in the small-ϵ limit. Note that the power-series expansion fails when $\delta = 1/2$; however, Equation (5) is sufficient to derive the limiting behavior as $\epsilon \to 0$ in the critical case.

Naively, Equation (8) can be inferred by directly substituting $\epsilon = 0$ into Equation (5). However, it is instructive to expand the quadratic equation in power series of ϵ, as in Equation (7), to obtain the limiting behavior. In fact, it is instrumental in deriving a nontrivial asymptotic mapping of parameters between the focal-link and focal-triad models (Section 3.2.3).

Subsequently, we can obtain approximate expressions for all r_k, based on the hierarchy of relaxation times. More specifically, r_3 (hence r_4 and r_5) can be identified using r_1 (Equation (6)); similarly, r_6 (hence r_7, r_8, and r_9) can be identified using r_3, r_4, and r_5. In practice, however, obtaining r_3 requires solving a cubic equation whose solution is complicated and therefore anything but useful to write down in explicit form. Thus, we use symbolic computations in PYTHON to determine all remaining rescaled frequencies, from r_3 through r_9 (see Supplemenatry Materials for details). Combining our mean-field analysis and computational results, we find that the equilibrium triad frequencies r_k in the activation-limited regime depend upon parameters in the following way:

$$\begin{cases} r_1 = r_2 = O(\epsilon^{\alpha}), \\ r_3 = r_5 = O(\epsilon^{2\alpha}), \quad r_4 = O(\epsilon^{2\alpha}), \\ r_6 = r_9 = O(\epsilon^{3\alpha}), \quad r_7 = r_8 = O(\epsilon^{3\alpha}), \end{cases} \tag{9}$$

where the exponent α changes as

$$\alpha(\delta) = \begin{cases} 1 & \delta \approx 0, \\ 1/2 & \delta = 1/2, \\ 0 & \delta \approx 1, \end{cases} \tag{10}$$

and the value of α drops sharply around $\delta = 1/2$ (see Figure 6). In fact, the δ-dependence (Equation (10)) exhibits a crossover phenomenon in the power-law ϵ-scaling, when the exponent suddenly drops from 1 to $1/2$ as δ approaches $1/2$ from below. The drastic change in the power-law exponent α around $\delta = 1/2$ in all variables strongly signifies the phase transition at $\delta = 1/2$, which is already implied by Equation (8). We will demonstrate later that this transition actually corresponds to the emergence of a giant component connected by active links. We finally note that the ϵ-scaling identified in Equation (9) is consistent with the assumption that the frequencies of triads located lower in Figure 1 can be disregarded compared to those located higher, confirming post-hoc that our approach is self-consistent in the activation-limited regime.

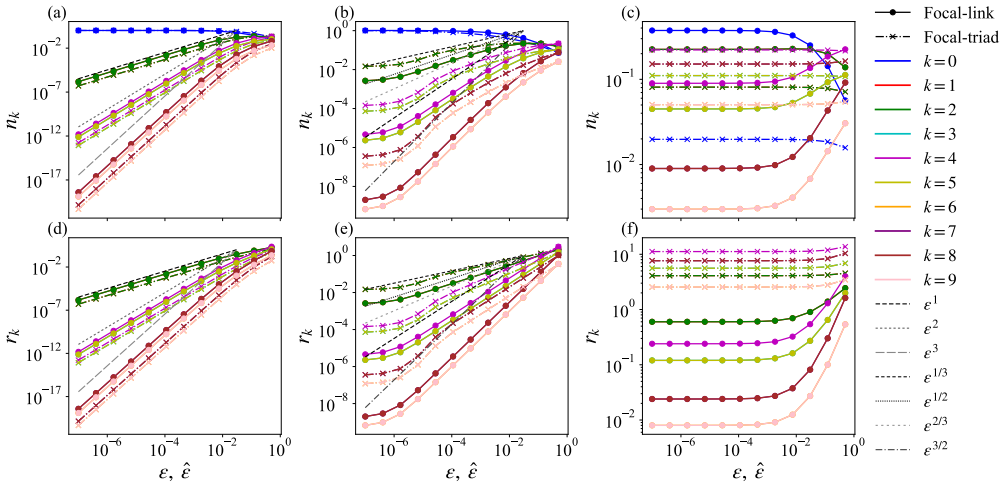

Figure 6. ϵ- or $\hat{\epsilon}$- scaling of r_k and n_k at equilibrium, obtained numerically from solving the rate equations in the subcritical (**a**,**d**), critical (**b**,**e**), and supercritical (**c**,**f**) regimes. Parameters: $\delta = 0.4$, $\eta = 0.5$ (focal-link model, (**a**,**d**)),

$\delta = 0.5$, $\eta = 0.5$ (focal-link model, (**b**,**e**)), $\delta = 0.6$, $\eta = 0.5$ (focal-link model, (**c**,**f**)), $\kappa_1 = 0.4$, $\kappa_2 = 0.6$, $\hat{\eta} = 0.5$ (focal-triad model, (**a**,**d**)), $\kappa_1 = 0.5$, $\kappa_2 = 0.5$, $\hat{\eta} = 0.5$ (focal-triad model, (**b**,**e**)), and $\kappa_1 = 0.6$, $\kappa_2 = 0.4$, $\hat{\eta} = 0.5$ (focal-triad model, (**c**,**f**)). These power-law behaviors can be consistently expressed in the hierarchical forms described by Equations (9) and (15). Some triads have nearly identical frequencies, causing some lines to overlap, as implied by Equation (1).

3.2.2. Focal-Triad Model: Scaling Behavior

We can likewise identify equilibrium rescaled frequencies in the focal-triad model. Using the same method, we find that r_1 approximately satisfies the following quadratic equation (see Supplemenatry Materials):

$$4\kappa_2 r_1^2 + 3(\kappa_2 - \kappa_1) r_1 - 3\hat{\epsilon} = 0, \tag{11}$$

whose positive solution is

$$r_1 = \frac{6\hat{\epsilon}}{3(\kappa_2 - \kappa_1) + \sqrt{9(\kappa_2 - \kappa_1)^2 + 48\kappa_2\hat{\epsilon}}}. \tag{12}$$

We also expand r_1 in powers of $\hat{\epsilon}$ (see Supplemenatry Materials for a derivation), giving

$$r_1 = \begin{cases} \frac{1}{\kappa_2 - \kappa_1}\hat{\epsilon} + O(\hat{\epsilon}^2) & \kappa_1 < \kappa_2, \\ \frac{3(\kappa_2 - \kappa_1)}{4\kappa_2} + \frac{1}{\kappa_1 - \kappa_2}\hat{\epsilon} + O(\hat{\epsilon}^2) & \kappa_1 > \kappa_2. \end{cases} \tag{13}$$

Equation (13) implies that a phase transition similar to the one identified in the focal-link model arises when $\kappa_1 = \kappa_2$ (provided $\kappa_2 \neq 0$):

$$r_1 \to \begin{cases} 0 & \kappa_1 \leq \kappa_2, \\ \frac{3(\kappa_1 - \kappa_2)}{4\kappa_2} & \kappa_1 \geq \kappa_2. \end{cases} \tag{14}$$

Further numerical results show the following scaling behavior with respect to parameters, clearly reminiscent of Equation (9):

$$\begin{cases} r_1 = r_2 = O(\hat{\epsilon}^\beta), \\ r_3 = r_5 = O(\hat{\epsilon}^{2\beta}), \quad r_4 = O(\hat{\epsilon}^{2\beta}), \\ r_6 = r_9 = O(\hat{\epsilon}^{3\beta}), \quad r_7 = r_8 = O(\hat{\epsilon}^{3\beta}), \end{cases} \tag{15}$$

where

$$\beta(\delta) = \begin{cases} 1 & \kappa_1 \ll \kappa_2, \\ 1/3 & \kappa_1 = \kappa_2, \\ 0 & \kappa_1 \gg \kappa_2, \end{cases} \tag{16}$$

and β changes drastically when $\kappa_1 = \kappa_2$ (see Figure 6). We note that the derived quadratic equation does not predict the critical exponent $\beta = 1/3$ (e.g., Equation (12) implies that r_1 increases in the order of $(\hat{\epsilon})^{1/2}$ when $\kappa_1 = \kappa_2$, which is inconsistent with the numerical results in Figure 6).

Finally, we note that Equations (9) and (10) (or Equations (15) and (16)) demonstrate that the results are consistent with the assumption that equilibrium triad frequencies are hierarchically ordered, if $\delta \leq 1/2$ (or $\hat{\kappa} \leq 1/2$). This implies that the activation-limited approximation maintains self-consistency within the subcritical and critical regimes.

3.2.3. Asymptotic Mapping between the Focal-Link and Focal-Triad Models

Although developed from different philosophies, the strikingly similar scaling behaviors between the focal-link and focal-triad models suggest an underlying connection between them. In fact, comparing the asymptotic behaviors captured by Equations (7) and (13) suggests the following mapping between δ and κ_1, κ_2 in the limit $\epsilon, \hat{\epsilon} \to 0$:

$$\delta \sim \hat{\kappa}(\kappa_1, \kappa_2), \tag{17}$$

where

$$\hat{\kappa}(\kappa_1, \kappa_2) = \frac{1}{2} + \frac{\kappa_1 - \kappa_2}{2}. \tag{18}$$

Note that we have identified the above mapping by comparing Equations (7) and (13) in the subcritical regime since the supercritical regime does not align with the activation-limited regime. In the same spirit, though less significantly, $\hat{\epsilon} \sim 3\epsilon$ is also obtained, constraining the relative rate at which ϵ and $\hat{\epsilon}$ tend to zero.

Equation (17) may be crucial because δ determines the strength of *signed* preferential reassessment, whereas $\hat{\kappa}$ facilitates social contagion or isolation, which activates or deactivates links *regardless* of the signs. This perhaps implies that since link signs, once activated, are constantly getting "mixed," it does not matter whether or not focal and neighboring links exhibit sign correlation via single updates when it comes to equilibrium triad frequencies because such one-time correlation vanishes in the long run. In short, this mapping suggests the strength of preferential reassessment is analogous to the balance between social contagion and social isolation. To examine if the asymptotic mapping works, we compared time-series data obtained from solving the rate equations corresponding to each model, given that $\delta = \hat{\kappa}$ and $\hat{\epsilon} = 3\epsilon$ are satisfied, and that ϵ is small. Figure 7 shows that equilibrium values are in excellent agreement between these two models, suggesting that they can indeed be mapped to each other via $\hat{\kappa}$. Note that this only works up to the critical point; equilibrium triad frequencies do not align beyond the subcritical regime.

The hallmark of the focal-triad model is the absence of *intentional* irrational updates, meaning that the focal triad never undergoes an irrational update event. However, irrational events can still happen *incidentally*: when a link flips its sign from positive to negative and vice versa, irrational updates may occur in a non-focal triad in which that link is also involved. Therefore, the actual number of irrational updates, whether intentional or incidental, is contingent on the equilibrium triad frequencies of each model. In Supplemenatry Materials Section S5, we show that the focal-triad model indeed substantially reduces the chance of such events in the activation-limited (or subcritical) regime. Note that this does not extend to the regime where most links are active, because irrational updates rarely occur in that scenario.

We have so far discussed how these two models are comparable, especially within the activation-limited regime. This correspondence extends to the supercritical regime, where a majority of links are active. Figure 8 shows how the equilibrium triad frequencies change with respect to δ (or $\hat{\kappa}$), focusing on the following four triad categories: inactive, semi-active, balanced, and unbalanced. Notably, Figure 8c shows that when η (or $\hat{\eta}$) is close to one, the semi-active triad frequencies increase with δ (or $\hat{\kappa}$) in the subcritical regime, and they decrease after δ (or $\hat{\kappa}$) surpasses the critical value. By contrast, all other triad frequencies—those of inactive, balanced, and unbalanced triads—change monotonically, either decreasing or increasing with δ (or $\hat{\kappa}$), across all panels. Again, these results suggest that although the supercritical regime is beyond the asymptotic mapping we have identified earlier, the two models exhibit similar behavior across all investigated regimes.

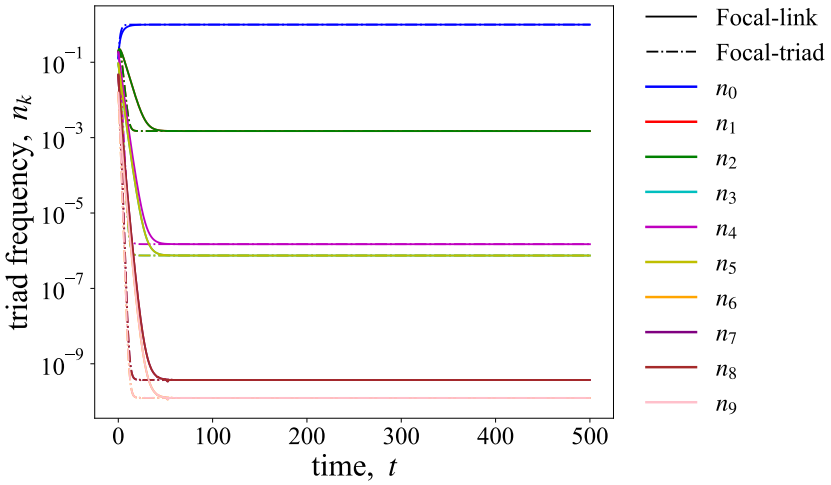

Figure 7. Time-series data compared between the focal-link and focal-triad models under the identified mapping (colors as in Figure 5). The equilibrium frequencies demonstrate an excellent agreement between them. Parameters: $\epsilon = 0.0001$, $\hat{\epsilon} = 0.0003$, $\delta = \kappa_1 = 0.4$, $\kappa_2 = 0.6$, $\eta = \hat{\eta} = 0.5$. The above choice of parameter values satisfies both $\hat{\kappa} = \delta$ and $\hat{\epsilon} = 3\epsilon$. Note that some triads have nearly identical frequencies, causing some lines to overlap, as implied by Equation (1).

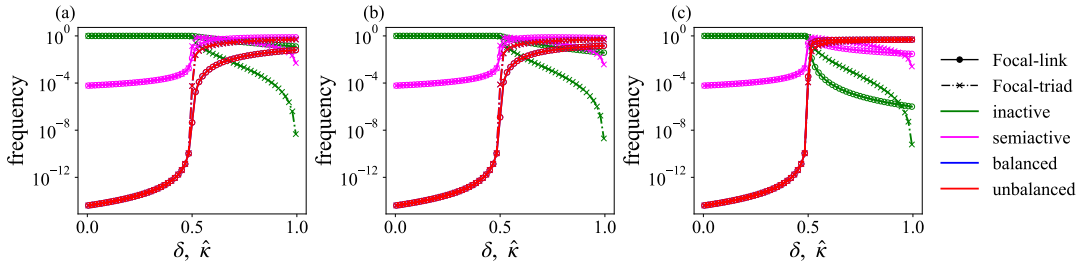

Figure 8. Equilibrium frequencies of inactive (three neutral links, n_0), semi-active (including one or two neutral links, $n_1 + n_2 + n_3 + n_4 + n_5$), balanced ($n_6 + n_8$), and imbalanced ($n_7 + n_9$) triads, obtained by numerically evolving the rate equations for a sufficiently long period of time. Parameters: (**a**) $\eta = \hat{\eta} = 0.01$, (**b**) $\eta = \hat{\eta} = 0.5$, and (**c**) $\eta = \hat{\eta} = 0.99$. Notably, the effect of δ and $\hat{\kappa}$ on the aggregated frequencies of the semi-active triads (Δ_1 through Δ_5) becomes non-monotonic, when η and $\hat{\eta}$ are close to one, respectively. Note that in the focal-triad model, the equilibrium frequencies of semi-active triads are non-monotonic in $\hat{\kappa}$ for all $\hat{\eta}$, whereas in the focal-link model, the non-monotonicity appears and becomes pronounced as η increases to one. In contrast, the frequencies of inactive, balanced, and imbalanced triads either increase or decrease monotonically with δ (or $\hat{\kappa}$), regardless of the values of η (or $\hat{\eta}$). Balanced and imbalanced triads are almost identically frequent across all these parameter regimes, making the blue lines and circles overlaid with the red ones. We set $\epsilon = 10^{-5}$ in the focal-link model, whereas in the focal-triad model, $\hat{\epsilon} = 3 \times 10^{-5}$, $\kappa_1 = \delta$, $\kappa_2 = 1 - \kappa_1$ so that $\hat{\kappa} = \delta$. To avoid numerical instability, we imposed $\delta, \hat{\kappa} < 1$ in all panels.

Nonetheless, an important difference between these two models arises in the supercritical regime. In the focal-link model, the non-monotonic effect of δ is pronounced only when η is close to 1, whereas in the focal-triad model, such effect is present regardless of the value of $\hat{\eta}$. The origin of this difference is not yet fully understood. In Supplemenatry Materials Figure S3, we also show how the equilibrium triad frequencies change when both δ and η (or $\hat{\kappa}$ and $\hat{\eta}$) change continuously, given a small fixed ϵ (or $\hat{\epsilon}$).

3.3. Intuitions for δ- and $\hat{\kappa}$-Transitions

We have analyzed the transition at $\delta = 1/2$ (or equivalently $\hat{\kappa} = 1/2$) using the mean-field approximation. Here we provide a simple intuition for why and how it occurs (see Figure 9 for an illustration). Let us begin with the focal-link model. Pick a Δ_1 generated through an ϵ-event that occurs with a small probability. Then, consider updating one of its inactive links through a δ-event. When δ is fixed at some value, the qualitative behavior of the system changes at $\delta = 1/2$. If $\delta > 1/2$, the likely outcome is that $O(N)$ Δ_1 triads are generated (as well as $O(1)$ Δ_3) per δ-event because most triads should be of type Δ_0 at the time of the update event, provided that $\epsilon \ll 1/2$. This also applies for Δ_2. Since n_1 and n_2 increase on the order of N every time the above event occurs, the top row triads Δ_1 and Δ_2 as well as Δ_0 quickly disappear; hence, the process is exponential. If $\delta < 1/2$, on the other hand, the likely outcome is that $O(1)$ Δ_1 and Δ_2 triads are generated per δ-event, which does not increase but rather maintains their overall frequencies (see Figure 9a).

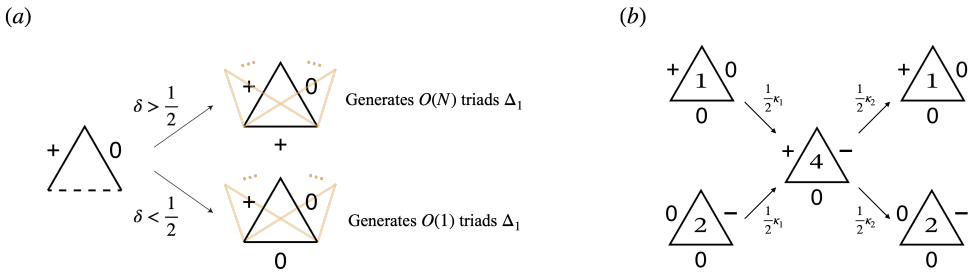

Figure 9. Illustration of the intuitions for the transition at $\delta = 1/2$ in (**a**) the focal-link model, and $\hat{\kappa} = 1/2$ in (**b**) the focal-triad model.

In the focal-triad model, the intuition is slightly different, due to structural differences underlying these models. A key observation in this case is the probability influx between adjacent triads in the diagram (Figure 4), and the overall behavior of the system is primarily governed by the flow between Δ_1 and Δ_4 and between Δ_2 and Δ_4 (see Figure 9b). When social contagion is equally as strong as social isolation, i.e., $\kappa_1 = \kappa_2$, the *upstream* and *downstream* flows balance out exactly between those two rows in Figure 1. This is because, given $\hat{\epsilon} \ll 1/2$, if the downward flow becomes more dominant than the upward one, i.e., $\kappa_1 > \kappa_2$, $O(N)$ Δ_1 (Δ_2) triads, in addition to $O(1)$ Δ_4 triads, will be generated from a Δ_2 (Δ_1) triad, much more than a Δ_4 triad generates $O(N)$ Δ_1 or Δ_2

triads. Since this process is recursive, the frequency of inactive links diminishes, again, exponentially. If κ_1 is less than κ_2, the instance of Δ_1 (or Δ_2) turning into Δ_4 is less likely to occur than the other way around. Therefore, Δ_1 or Δ_2 triads generated by $\hat{\epsilon}$-events do not lead to an exponential, chain reaction that generates $O(N)$ Δ_4 triads. Note that once most triads are of type Δ_3, Δ_4, or Δ_5, the lowest downstream flows, the so-called social balance effect, causes them to transition to bottom row triads, give that $\kappa_1 > \kappa_2$.

3.4. Largest Connected Component

A more intuitive way to understand the dynamics of our models is to focus on active links, instead of ten different types of triads, and to consider them as social networks evolving via triadic interactions. One of the fundamental quantities that characterize static social networks is the size of the largest connected component. In network science, a *weakly connected component* of an undirected graph refers to a set of nodes in which any two of them are connected through one or more undirected edges in that graph [32]. To avoid unnecessary wordiness, however, we refer to it as a *connected component* hereafter, and we are primarily interested in the size of the *largest* connected component. When it spans most of the network, it is called *a giant connected component*. More precisely, a giant connected component is a connected component whose size grows proportional to the number of nodes [32]. Here we focus on the largest *positively* connected component, i.e., the largest component connected by positive links, so that the largest connected component size is meaningful in social contexts and therefore provides more plausible insight.

To see how this quantity changes with parameters, we measured the size of the largest positively connected component averaged over a sufficiently long period of time, for different values of ϵ and δ (or $\hat{\epsilon}$ and $\hat{\kappa}$; see Figure 10). In both models, the largest connected component size, denoted $\mathcal{S}$, dramatically increases at $\delta = 1/2$ ($\hat{\kappa} = 1/2$) for small ϵ ($\hat{\epsilon}$)—roughly from $O(1)$ to $O(N)$. This implies that the phase transition we have identified in our mean-field analysis corresponds to the emergence of a giant connected component in the population. On the other hand, along the ϵ-axis ($\hat{\epsilon}$-axis), the decrease in the size of the largest connected component is quite gradual until the parameter becomes very small; in other words, $\mathcal{S}$ approaches zero increasingly as $\epsilon \to 0$.

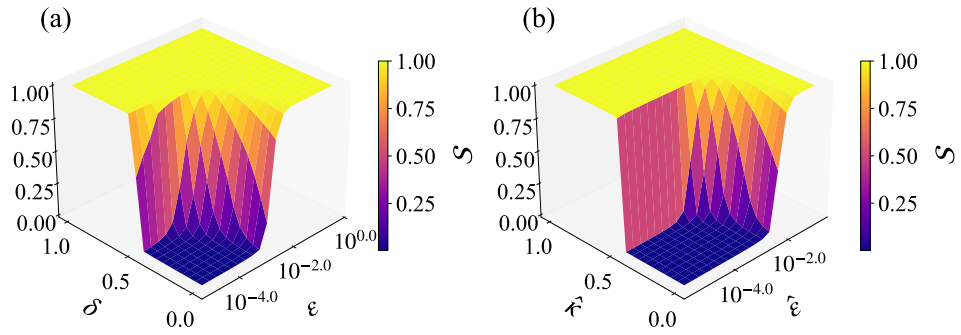

Figure 10. The largest connected component size in networks of size $N = 512$, obtained by time-averaging the data over the last half of 2000 Monte Carlo simulation steps, and rescaled by the network size N, in (**a**) the focal-link model and (**b**) the focal-triad model. The ϵ- and $\hat{\epsilon}$-axes are in logarithmic scale. Both the vertical axis and colorbar represent the size of the largest connected component. A giant (positively) connected component emerges at $\delta = 1/2$ and $\hat{\kappa} = 1/2$. This corresponds to the phase transition identified using the mean-field analysis in Section 3.2. Parameters: $\eta = 0.5$, $\hat{\eta} = 0.5$.

3.5. The Critical Regime: Connected but Not Demanding

We have identified the emergence of a giant connected component in both models by time-averaging simulation data, and it is well-predicted by the mean-field analysis as well as more intuitive arguments. However, a key question remains: what does the emergence of a giant connected component imply in the context of social networks and human behavior? How is it relevant to any real social systems? Here we provide a simple insight regarding why the critical regime is the only scenario that is both attainable and functional in real social situations.

We focus this discussion on the focal-link model. As ϵ tends to zero, a phase transition occurs at $\delta = 1/2$ in the rescaled size of the largest connected component size, rapidly increasing from 0 to 1. We define individual *cognitive load* by the per-capita number of "activation events" during a unit time. By *activation event* we refer to any update event that changes the sign of a link from 0 to either 1 or -1. This means that two unacquainted or "mutually neutral" individuals become either friends with or enemies of each other. Getting to know someone requires a nontrivial amount of time and effort, imposing some natural limit to the number of well-developed mutually friendly

(or unfriendly) relationships an individual can maintain. In other words, activating a link is costly [30]. We explore how much per-capita cognitive load is required in our framework, particularly in the large-N limit, provided the equilibrium triad frequencies identified in Section 3.1.

Let $\mathcal{L}$ be the mean individual cognitive load per unit time. Counting all instances of activating inactive links and dividing the number by population size N lead to

$$\mathcal{L} = 2\frac{L}{N}\left(p_0^+ + p_0^-\right) \tag{19}$$

where $L = \binom{N}{2}$ represents the total number of links and ρ^0 is the density of neutral links (see Supplemenatry Materials for notations and expressions for $p_\sigma^{\sigma'}$). The prefactor accounts for the fact that each link update involves two individuals. Assuming that $\epsilon \ll 1/2$, we substitute Equation (9) into Equation (19) to find an approximate formula for $\mathcal{L}$. When $\delta = 1/2$, $\mathcal{L}$ scales as $O(\epsilon N)$, which means that setting $\epsilon = O(N)$ ensures that $\mathcal{L}$ be bounded even when the population size N increases. When $\delta > 1/2$, on the other hand, $\mathcal{L}$ grows as $O(N)$, which becomes unbounded as N tends to infinity, regardless of how we set ϵ or any other parameters. When $\delta < 1/2$, however, the expected size of the largest connected component size is $O(1)$, according to Figure 10. Therefore, the critical regime is the only scenario where individual cognitive load can be bounded—when ϵ is set to decrease on the order of $1/N$—*and* a giant connected component can still emerge in the limit of a large population size. A similar argument applies for the focal-triad model. This suggests that only the regime where $\delta = 1/2$ and $\epsilon = O(N^{-1})$ produces a (constantly evolving) society that can prevent individual cognitive load from increasing limitlessly while still achieving decent inter-individual connectivity in a growing population.

4. Discussion

We have developed and analyzed two dynamic models of social balance where interpersonal relationships assume three possible states. One of our key findings is that, contrary to the classical case with a finite population, polarization does not occur in these extended models, except in measure-zero cases that correspond to the classical scenario (see also Supplemenatry Materials Figure S5, which shows that initial complete balance can collapse into transient dynamics in near-classical yet distinct settings). Mean-field theory successfully predicts equilibrium triad frequencies in both models, which is not the case for traditional models with a finite population where interpersonal relationship states are binary. There is also an asymptotic mapping between these two alternative models, admitting one-to-one correspondence at the level of equilibrium triad frequencies. Subsequently, we have identified the emergence of a giant component connected by active links when a (normalized) parameter exceeds a critical value. Using our mean-field analysis, we have also interpreted the critical regime in terms of per-capita cognitive load: the critical regime can both achieve nontrivial inter-individual connectedness and keep individual cognitive load bounded at the same time, even as the population size grows infinitely.

When all the (normalized) parameters are set to their maximum values, both of our models recover the classical scenario, in which case they eventually reach fully balanced states (paradise or full polarization) regardless of initial conditions, in agreement with previous studies. Otherwise, however, transient dynamics dominate: the population almost never reaches a fully polarized or paradise state, and even though triad frequencies reach stable equilibria, any given link between a pair of individuals continues to stochastically switch signs over—that is, a dynamic steady state. Therefore, although it is natural to extend traditional models of social balance to accommodate neutral links, this extension can completely alter the long-term dynamical outcomes. This basic finding hints at something more fundamental to balance theory: the binary representation of interpersonal relationships may underlie societal polarization caused via social balance. When divided societies occur, it may be due to significantly strong inter-individual connectivity as well as the tendency to simplify situations, such as favoring friends or enemies over more nuanced, complex, and yet, sometimes, unsettling relationships between two individuals.

Neutrality may alternatively be seen as a key to preventing polarization in a society. Today, many societies are suffering from deep and seemingly irreversible polarization [33–37]. The recent surge of studies on polarization underscores the urgency to pin down its causes and consequences, and to identify effective measures to preempt or counteract persistent partisanship [38–41]. Previous studies have proposed several key factors, such as wealth redistribution [42] and attraction to self-oriented ideologies [43]. Our study shows that neutral ties—and dynamical updates that account for neutral ties—can also prevent polarization in the realm of balance theory, as prior empirical work has suggested [30,44].

One of the salient features of our models is that the emergence of a giant connected component is characterized by a phase transition. The critical regime in fact suggests that individual cognitive load at equilibrium remains bounded, yet a large connected society could emerge. On the one hand, it should hardly be considered a society

when people are isolated and barely engage in groups. On the other hand, humans can maintain only a limited number of meaningful connections—also known as Dunbar's number. When properly adjusted, the tradeoff between these two factors results in a sweet spot—the critical regime—that can achieve the right balance between them, leading to sustainable societies. Some complex systems even exhibit self-organized criticality [45,46]. Although whether this applies to social systems remains an active area of study, our results suggest a clear advantage of being close to the critical regime within the frameworks of balance theory with ternary interpersonal relationships.

One possible application of this study is to compare the equilibrium proportions of different triad configurations identified in our model to those in real-world social networks, at least in terms of the orders of magnitude of those frequencies. Prior empirical studies have analyzed signed networks involving neutral ties, including in online, social, and political settings [29,30]. Extending those data and approaches would help us to understand real-world networks in the context of neutral links. An alternative but coarser approach would be to compare our models to data at the level of network characteristics. Emergent network properties and structures in the critical regime in our models could be explored further, as some nontrivial network properties have been associated with criticality and documented in some empirical social networks [3,45,46]. This may reveal how triadic interactions govern the way social networks form and reorganize themselves over time.

There is substantial overlap between indirect reciprocity and social balance at a conceptual level. In simple terms, indirect reciprocity is a mechanism for the emergence of cooperation by conditioning behavior on the reputations of others and updating reputations based on their behavior toward third parties [47–51]. Traag et al. have extended the notion of balance to directed interpersonal relationships, i.e., gossiping, and have found that different mechanisms of gossip lead to balance or imbalance, where homophilic interactions are better suited for achieving cooperation [52]. Oishi et al. have also developed a model that bridges social balance and indirect reciprocity where people can have different opinions of each other (i.e., private assessment); they found that a balanced (frozen) state can only be achieved when the underlying network is almost fully connected if not too sparse [53]. In retrospect, those results suggested a fundamental character of social balance theory: social balance is not robust to structural changes in the models. Our models and analysis derive the same conclusion: when relationships are ternary, networks are not fully connected, and consequently dynamics are more intricate, making globally balanced states unlikely to occur.

Both of our models have their own shortcomings, which are partly compensated by each other. In the focal-link model, the current status of a link is ignored during an update, which seems unrealistic in some social settings. In the focal-triad model, the parameter that governs the transitions between the bottom-row triads has little impact on the overall equilibrium triad frequencies. Although we have focused our analysis on equilibrium triad densities and network connectivity, exploring network structures could be an interesting avenue for future research, including degree distributions, network topologies, and connected component size distributions.

Our models can be extended in several directions. First, each update currently concerns only three individuals. While this aligns with the dynamical model introduced in Ref. [25], called the *local triad dynamics*, it contrasts the other model introduced in the same paper called the *constrained triad dynamics*, in which every update considers all the triads that involve a focal link to be updated. Given that the constrained dynamics promotes overall balance more effectively, taking into account *all* the triads that are attached to a focal link may contribute more to polarized outcomes (i.e., full balance) than considering only a single triad. Flipping multiple links in a single update may also change the results significantly, as this would allow for instantaneous transitions between triads that do not directly correspond via a single-link flip and thus are currently prohibited. Extending our models in this direction would require some nontrivial reformulation; and it may require empirical study on whether simultaneous multi-link updates are plausible, especially since this possibility has not been addressed in prior models of social balance even in binary settings [16,25,51,53,54]. Furthermore, individual intrinsic differences in behavior, attitude, or belief are missing in all these models, which could possibly account for more realistic phenomena, such as *localized* polarization. All of these directions for future work will enrich our understanding of a comprehensive theory of social balance, especially in the presence of neutral links.

Supplementary Materials

The additional data and information can be downloaded at: https://media.sciltp.com/articles/others/2512181530131692/J.Soc.Phys.-25100083-Supplementary-Materials.pdf.

Author Contributions

H.G., M.S., H.N. and J.B.P. designed research; H.G. performed research; H.G., M.S., H.N. and J.B.P. analyzed data; H.G. and J.B.P. wrote the paper with input from M.S. and H.N. All authors have read and agreed to the published version of the manuscript.

Funding

H.G. acknowledges support from the Japan Society for the Promotion of Science (JSPS) under the Grant-in-Aid for JSPS Fellows No. JP25KJ2129, from the Japan Science Society under the Sasakawa Scientific Research Grant No. 2024-6036, and from the Meiji University Overseas Challenge Program. M.S. acknowledges support from JSPS KAKENHI Grant No. JP22K17971. M.S. and H.N. acknowledge support from JSPS KAKENHI Grant No. JP21H05297. J.B.P. acknowledges support from the US Army Research Office (award W911NF2410393) and the US Office of Naval Research (award N000142412778).

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

No AI tools were utilized for this paper.

References

1. Wasserman, S.; Faust, K. *Social Network Analysis*; Cambridge University Press: Cambridge, UK, 1994.
2. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **2009**, *81*, 591–646.
3. Holme, P.; Newman, M.E.J. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev.* **2006**, *74*, 56108.
4. Marvel, S.A.; Hong, H.; Papush, A.; et al. Encouraging Moderation: Clues from a Simple Model of Ideological Conflict. *Phys. Rev. Lett.* **2012**, *109*, 118702.
5. Baumann, F.; Lorenz-Spreen, P.; Sokolov, I.M.; et al. Modeling Echo Chambers and Polarization Dynamics in Social Networks. *Phys. Rev. Lett.* **2020**, *124*, 48301.
6. Mittal, D.; Constantino, S.M.; Vasconcelos, V.V. Anticonformists catalyze societal transitions and facilitate the expression of evolving preferences. *PNAS Nexus* **2024**, *3*, 302.
7. Zachary, W.W. An Information Flow Model for Conflict and Fission in Small Groups. *J. Anthropol. Res.* **1977**, *33*, 452–473.
8. Facchetti, G.; Iacono, G.; Altafini, C. Computing global structural balance in large-scale signed social networks. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20953–20958.
9. Szell, M.; Lambiotte, R.; Thurner, S. Multirelational organization of large-scale social networks in an online world. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 13636–13641.
10. Belaza, A.M.; Hoefman, K.; Ryckebusch, J.; et al. Statistical physics of balance theory. *PloS ONE* **2017**, *12*, e0183696.
11. Pham, T.M.; Korbel, J.; Hanel, R.; et al. Empirical social triad statistics can be explained with dyadic homophylic interactions. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2121103119.
12. Galesic, M.; Olsson, H.; Pham, T.M.; et al. Experimental evidence confirms that triadic social balance can be achieved through dyadic interactions. *NPJ Complex.* **2025**, *2*, 1.
13. Leskovec, J.; Huttenlocher, D.; Kleinberg, J. Signed networks in social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 1361–1370.
14. Estrada, E.; Benzi, M. Walk-based measure of balance in signed networks: Detecting lack of balance in social networks. *Phys. Rev.* **2014**, *90*, 042802.
15. Doreian, P.; Mrvar, A. A partitioning approach to structural balance. *Soc. Netw.* **1996**, *18*, 149–168.
16. Antal, T.; Krapivsky, P.L.; Redner, S. Social balance on networks: The dynamics of friendship and enmity. *Phys. Nonlinear Phenom.* **2006**, *224*, 130–136.
17. Lee, J.H.; Sato, N.; Yano, K.; et al. Universal association between depressive symptoms and social-network structures in the workplace. *Sci. Rep.* **2022**, *12*, 10170.
18. Auster, C.J. Balance Theory and other Extra-Balance Properties: An Application to Fairy Tales. *Psychol. Rep.* **1980**, *47*, 183–188.

19. Saiz, H.; Gómez-Gardeñes, J.; Nuche, P.; et al. Evidence of structural balance in spatial ecological networks. *Ecography* **2017**, *40*, 733–741.
20. Ilany, A.; Booms, A.S.; Holekamp, K.E. Topological effects of network structure on long-term social network dynamics in a wild mammal. *Ecol. Lett.* **2015**, *18*, 687–695.
21. Rötzmeier-Keuper, J.; Hendricks, J.; Wünderlich, N.V.; et al. Triadic relationships in the context of services for animal companions. *J. Bus. Res.* **2018**, *85*, 295–303.
22. Heider, F. Attitudes and Cognitive Organization. *J. Psychol.* **1946**, *21*, 107–112.
23. Cartwright, D.; Harary, F. Structural balance: a generalization of Heider's theory. *Psychol. Rev.* **1956**, *63*, 277–293.
24. Harary, F. On the notion of balance of a signed graph. *Mich. Math. J.* **1953**, *2*, 143–146.
25. Antal, T.; Krapivsky, P.L.; Redner, S. Dynamics of social balance on networks. *Phys. Rev.* **2005**, *72*, 36121.
26. Kułakowski, K.; Gawroński, P.; Gronek, P. The Heider Balance: A Continuous Approach. *Int. J. Mod. Phys.* **2005**, *16*, 707–716.
27. Marvel, S.A.; Kleinberg, J.; Kleinberg, R.D.; Strogatz, S.H. Continuous-time model of structural balance. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1771–1776.
28. Goto, H.; Shiraishi, M.; Nishimori, H. Onset of Intragroup Conflict in a Generalized Model of Social Balance. *Phys. Rev. Lett.* **2024**, *133*, 127402.
29. Belaza, A.M.; Ryckebusch, J.; Bramson, A.; et al. Social stability and extended social balance—Quantifying the role of inactive links in social networks. *Phys. Stat. Mech. Its Appl.* **2019**, *518*, 270–284.
30. Dekker, D.; Krackhardt, D.; Doreian, P.; et al. Balance correlations, agentic zeros, and networks: The structure of 192 years of war and peace. *PLoS ONE* **2024**, *19*, e0315088.
31. White, H.C.; Boorman, S.A.; Breiger, R.L. Social Structure from Multiple Networks. I. Blockmodels of Roles and Positions. *Am. J. Sociol.* **1976**, *81*, 730–780.
32. Newman, M. *Networks*; Oxford University Press: Oxford, UK, 2018.
33. Baumann, F.; Lorenz-Spreen, P.; Sokolov, I.M.; et al. Emergence of Polarized Ideological Opinions in Multidimensional Topic Spaces. *Phys. Rev.* **2021**, *11*, 11012.
34. Waller, I.; Anderson, A. Quantifying social organization and political polarization in online platforms. *Nature* **2021**, *600*, 264–268.
35. Santos, F.P.; Lelkes, Y.; Levin, S.A. Link recommendation algorithms and dynamics of polarization in online social networks. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2102141118.
36. Liu, J.; Huang, S.; Aden, N.M.; et al. Emergence of Polarization in Coevolving Networks. *Phys. Rev. Lett.* **2023**, *130*, 37401.
37. Peralta, A.F.; Ramaciotti, P.; Kertész, J.; et al. Multidimensional political polarization in online social networks. *Phys. Rev. Res.* **2024**, *6*, 13170.
38. Levin, S.A.; Milner, H.V.; Perrings, C. The dynamics of political polarization. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2116950118.
39. Wang, S.S.H.; Cervas, J.; Grofman, B.; et al. A systems framework for remedying dysfunction in US democracy. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2102154118.
40. Bednar, J. Polarization, diversity, and democratic robustness. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2113843118.
41. Baldassarri, D.; Page, S.E. The emergence and perils of polarization. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2116863118.
42. Stewart, A.J.; Plotkin, J.B.; McCarty, N. Inequality, identity, and partisanship: How redistribution can stem the tide of mass polarization. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2102140118.
43. Axelrod, R.; Daymude, J.J.; Forrest, S. Preventing extreme polarization of political attitudes. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2102139118.
44. Lerner, J. Structural balance in signed networks: Separating the probability to interact from the tendency to fight. *Soc. Netw.* **2016**, *45*, 66–77.
45. Tadić, B.; Dankulov, M.M.; Melnik, R. Mechanisms of self-organized criticality in social processes of knowledge creation. *Phys. Rev.* **2017**, *96*, 32307.
46. Muñoz, M.A. *Colloquium*: Criticality and dynamical scaling in living systems. *Rev. Mod. Phys.* **2018**, *90*, 31001.
47. Nowak, M.A.; Sigmund, K. Evolution of indirect reciprocity. *Nature* **2005**, *437*, 1291–1298.
48. Sigmund, K. Moral assessment in indirect reciprocity. *J. Theor. Biol.* **2012**, *299*, 25–30.
49. Tanabe, S.; Suzuki, H.; Masuda, N. Indirect reciprocity with trinary reputations. *J. Theor. Biol.* **2013**, *317*, 338–347.
50. Murase, Y.; Kim, M.; Baek, S.K. Social norms in indirect reciprocity with ternary reputations. *Sci. Rep.* **2022**, *12*, 455.
51. Bae, M.; Shimada, T.; Baek, S.K. Indirect reciprocity as a dynamics for weak balance. *Phys. Rev.* **2025**, *112*, L032304.
52. Traag, V.A.; Dooren, P.V.; Leenheer, P.D. Dynamical Models Explaining Social Balance and Evolution of Cooperation. *PLoS ONE* **2013**, *8*, e60063.
53. Oishi, K.; Miyano, S.; Kaski, K.; Shimada, T. Balanced-imbalanced transitions in indirect reciprocity dynamics on networks. *Phys. Rev.* **2021**, *104*, 24310.
54. Pham, T.M.; Kondor, I.; Hanel, R.; Thurner, S. The effect of social balance on social fragmentation. *J. R. Soc. Interface* **2020**, *17*, 20200752.