Logit Dynamics with Concurrent Updates for Local-Interaction Games*

Vincenzo Auletta† Diodato Ferraioli‡ Francesco Pasquale§ Paolo Penna¶
Giuseppe Persiano‖

Abstract

Logit choice dynamics are a family of randomized best response dynamics based on the *logit choice function* [28] that are used to model players with limited rationality and knowledge. In this paper we study the *all-logit dynamics*, where at *each* time step *all* players *concurrently* update their strategies according to the logit choice function. In the well studied *one-logit dynamics* [9] instead at each step *only one* randomly chosen player is allowed to update.

We study properties of the all-logit dynamics in the context of *local interaction games*, a class of games that has been used to model complex social phenomena [9, 31, 35] and physical systems [25]. In a local interaction game, players are the vertices of a *social graph* whose edges are two-player potential games. Each player picks one strategy to be played for all the games she is involved in and the payoff of the player is the (weighted) sum of the payoffs from each of the games. We prove that local interaction games characterize the class of games for which the all-logit dynamics is reversible.

We then compare the stationary behavior of one-logit and all-logit dynamics. Specifically, we look at the expected value of a notable class of observables, that we call *decomposable* observables. We prove that the difference between the expected values of the observables at stationarity for the two dynamics depends only on β (the *rationality* level) and on the distance of the social graph from a bipartite graph. In particular, if the social graph is bipartite then decomposable observables have the same expected value. Finally, we show that for some games the mixing times of one-logit and all-logit dynamics are almost equivalent.

---

*This work is partially supported by PRIN 2008 research project COGENT (COmputational and GamE-theoretic aspects of uncoordinated NeTworks), funded by the Italian Ministry of University and Research.

†Dipartimento di Informatica, Università di Salerno. Email: `auletta@dia.unisa.it`.

‡LAMSADE, Université Paris-Dauphine. Email: `diodato.ferraioli@dauphine.fr`. This author is supported by ANR, project COCA, ANR-09-JCJC-0066.

§Dipartimento Bo, University Bo. Email: `pasquale@bo.bo`.

¶Institute of Theoretical Computer Science, ETH Zurich. Email: `paolo.penna@inf.ethz.ch`.

‖Dipartimento di Informatica, Università di Salerno. Email: `giuper@dia.unisa.it`.

1 Introduction

In the last decade, we have observed an increasing interest in understanding phenomena occurring in complex systems consisting of a large number of simple networked components that operate autonomously guided by their own objectives and influenced by the behavior of the neighbors. Even though (online) social networks are a primary example of such systems, other remarkable typical instances can be found in Economics (e.g., markets), Physics (e.g., Ising model and spin systems) and Biology (e.g., evolution of life). A common feature of these systems is that the behavior of each component depends only on the interactions with a limited number of other components (its neighbors) and these interactions are usually very simple.

Game Theory is the main tool used to model the behavior of agents that are guided by their own objective in contexts where their gains depend also on the choices made by neighboring agents. Game theoretic approaches have been often proposed for modeling phenomena in a complex social network, such as the formation of the social network itself [21, 6, 3, 15, 12, 11, 10], and the formation of opinions [23, 8, 16] and the spread of innovation [34, 35, 31] in the social network. Many of these models are based on *local interaction games*, where agents are represented as vertices on a *social graph* and the relationship between two agents is represented by a simple two-player game played on the edge joining the corresponding vertices.

We are interested in the *dynamics* that governs such phenomena and several dynamics have been studied in the literature like, for example, the best response dynamics [18], the logit dynamics [9], fictitious play [17] or no-regret dynamics [20]. Any such dynamics can be seen as made of two components:

- *Selection rule:* by which the set of players that update their state (strategy) is determined;
- *Update rule:* by which the selected players update their strategy.

For example, the classical best response dynamics composes the *best response* update rule with a selection rule that selects one player at the time. In the best response update rule, the selected player picks the strategy that, given the current strategies of the other players, guarantees the highest utility. The Cournot dynamics [13] instead combines the best response update rule with the selection rule that selects all players. Other dynamics in which all players concurrently update their strategy are fictitious play [17] and the no-regret dynamics [20].

In this paper, we study a specific class of randomized update rules called the *logit choice function* [28, 9, 33] which is a type of noisy best response that models in a clean and tractable way the limited knowledge (or bounded rationality) of the players in terms of a parameter β called *inverse noise*. In similar models studied in Physics, β is the inverse of the temperature. Intuitively, a low value of β (that is, high temperature) models a noisy scenario in which players choose their strategies "nearly at random"; a high value of β (that is, low temperature) models a scenario with little noise in which players pick the strategies yielding higher payoffs with higher probability.

The logit choice function can be coupled with different selection rules so to give different dynamics. For example, in the *logit* dynamics [9] at every time step a single player is selected uniformly at random and the selected player updates her strategy according to the logit choice function. The remaining players are not allowed to revise their strategies in this time step. One of the appealing features of the logit dynamics is that it naturally describes an ergodic Markov chain. This means that the underlying Markov chain admits a *unique stationary distribution* which we take as solution concept. This distribution describes the long-run behavior of the system (which states appear more frequently over a long run). The interplay between the noise and the underlying game naturally determines the system behavior: (i) As the noise becomes "very large" the equilibrium point is "approximately" the uniform distribution; (ii) As the noise vanishes the stationary distribution concentrates on so called stochastically stable states which, for certain classes of games, correspond to pure Nash equilibria.

While the logit choice function is a very natural behavioral model for approximately rational agents, the specific selection rule that selects one single player per time step avoids any form of concurrency. Therefore a natural question arises

> What happens if *concurrent* updates are allowed?

For example, it is easy to construct games for which the best response converges to a Nash equilibrium when only one player is selected at each step and does not converge to any state when more players are chosen to concurrently update their strategies.

In this paper we study how the logit choice function behave in an extremal case of concurrency. Specifically, we couple this update rule with a selection rule by which *all* players update their strategies at every time step. We call such a dynamics *all-logit*, as opposed to the classical (*one-*)logit dynamics in which only one player at a time is allowed to move. Roughly speaking, the all-logit is to the one-logit what the Cournot dynamics is to the best response dynamics.

Our contributions. We study the all-logit dynamics for local interaction games [14, 31]. Here players are vertices of a graph, called the *social graph*, and each edge is a two-player game. We remark that games played on different edges by a player may be different but, nonetheless, they have the same strategy set for the player. Each player picks one strategy that is used for all of her edges and the payoff is a (weighted) sum of the payoffs obtained from each game. This class of games includes coordination games on a network [14] that have been used to model the spread of innovation and of new technology in social networks [34, 35], and the Ising model [27], a model for magnetism. In particular, we study the all-logit dynamics on local interaction games for every possible value of the inverse noise β and we are interested on properties of the original one-logit dynamics that are preserved by the all-logit.

As a warm-up, we discuss two classical two-player games (these are trivial local interaction games played on a graph with two vertices and one edge): the coordination game and the prisoner's dilemma. Even though for both games the stationary distribution of the one-logit and of the all-logit are quite different, we identify three similarities. First, for both games, both Markov chains are reversible. Moreover, for both games, the expected number of players playing a certain strategy at the stationarity of the all-logit is exactly the same as if the expectation was taken on the stationary distribution of the one-logit. Finally, for these games the mixing time is asymptotically the same regardless of the selection rule. In this paper we will show that none of these findings is accidental.

We first study the *reversibility* of the all-logit dynamics, an important property of stochastic processes that is useful also to obtain explicit formulas for the stationary distribution. We *characterize* the class of games for which the all-logit dynamics (that is, the Markov chain resulting from the all-logit dynamics) is reversible and it turns out that this class coincides with the class of local interaction games. This implies that the all-logit dynamics of all two-player potential games are reversible; whereas not all potential games have a reversible all-logit dynamics. This is to be compared with the well-known result saying that one-logit dynamics of every potential game is reversible with respect to the Gibbs measure [9]. One of the tools we develop for our characterization yields a closed formula for the stationary distribution of reversible all-logit dynamics.

Then, we focus on the *observables* of local interaction games. An observable is a function of the strategy profile (that is the sequence of strategies adopted by the players) and we are interested in its expected values at stationarity for both the one-logit and the all-logit. A prominent example of observable is the difference Diff between the number of players adopting two given strategies in a game. In a local interaction game modeling the spread of innovation on a social network this observable counts the difference between the number of adopters of the new and old technology whereas in the Ising model it is the magnetic field of a magnet.

We show that there exists a class of observables whose expectation at stationarity of the all-logit is the same as the expectation at stationarity of the one-logit as long as the social network underlying the local interaction game is bipartite (and thus trivially for all two-player games). This class of observables includes the Diff observable. It is interesting to note that the Ising game has been mainly studied for bipartite graphs (e.g., the two-dimensional and the three-dimensional lattice). This implies that, for the Ising model, the all-logit is a dynamics that is compatible with the observations and it is arguably more natural than the one-logit (that postulates that at any given time step only one particle updates its status and that the update strategy is instantaneously propagated). We extend this result by showing that for general graphs, the extent at which the expectations of these observables differ can be upper and lower bounded by a function of β and of the distance of the social graph from a bipartite graph.

Finally, we give the first bounds on the mixing time of the all-logit. We start by giving a *general* upper bound on the mixing time of the all-logit in terms of the *cumulative utility* of the game. We then look at a specific well-known n-player local interaction game: the Curie-Weiss model from Statistical Physics and derive an upper bound on the mixing time that is tighter than the one obtained from our general upper bound. We complement the upper bound with a lower bound. The two bounds show that, for n players, the mixing time is constant for $\beta = \mathcal{O}(1/n^2)$, polynomial for $\beta = \mathcal{O}(\log n/n^2)$, and exponential for $\beta = \Omega(1/n)$. The mixing time for β between $\log n/n^2$ and $1/n$ is still open. Asymptotically, these bounds almost match the ones known for the one-logit dynamics.

Related works on logit dynamics. The all-logit dynamics for strategic games has been studied by Alos-Ferrer and Netzer [1]. Specifically, in [1] the authors study the logit-choice function combined with general selection rules (including the selection rule of the all-logit) and investigate conditions for which a state is *stochastically stable*. A stochastically stable state is a state that has non-zero probability as β goes to infinity. We focus instead on a specific selection rule that is used by several remarkable dynamics considered in Game Theory (Cournot, fictitious play, and no-regret) and consider the whole range of values of β.

The one-logit dynamics has been actively studied starting from the work of Blume [9]that showed that for 2×2 coordination games, the risk dominant equilibria (see [19]) are stochastically stable. Much work has been devoted to the study of the one-logit for local interaction games with the aim of modeling and understanding the spread of innovation in a social network [14, 35]. A general upper bound on the mixing time of the one-logit dynamics for this class of games is given by Berger et al. [7]. Montanari and Saberi [31]instead studied the hitting time of the highest potential configuration and relate this quantity to a connectivity property of the underlying network. The mixing time and the metastability of the one-logit dynamics for strategic games have been studied in [4, 5].

2 Definitions

In this section we formally define the local interaction games and the Markov chain induced by the all-logit dynamics.

Strategic games. Let $\mathcal{G} = ([n], S_1, \ldots, S_n, u_1, \ldots, u_n)$ be a finite normal-form strategic game. The set $[n] = \{1, \ldots, n\}$ is the player set, S_i is the set of *strategies* for player $i \in [n]$, $S =$

$S_1 \times S_2 \times \cdots \times S_n$ is the set of *strategy profiles* and $u_i \colon S \to \mathbb{R}$ is the *utility* function of player $i \in [n]$.

We adopt the standard game-theoretic notation and denote by S_{-i} the set $S_{-i} = S_1 \times \ldots \times S_{i-1} \times S_{i+1} \times \ldots S_n$ and, for $\mathbf{x} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in S_{-i}$ and $y \in S_i$, we denote by $(\mathbf{x}, y)$ the strategy profile $(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \in S$.

Potential games [30] are an important class of games. We say that function $\Phi : S \to \mathbb{R}$ is an *exact potential* (or simply a *potential*) for game $\mathcal{G}$ if for every $i \in [n]$ and every $\mathbf{x} \in S_{-i}$

$$u_i(\mathbf{x}, y) - u_i(\mathbf{x}, z) = \Phi(\mathbf{x}, z) - \Phi(\mathbf{x}, y)$$

for all $y, z \in S_i$. A game $\mathcal{G}$ that admits a potential is called a *potential game.*

Local interaction games. In a *local interaction game* $\mathcal{G}$, each player i, with strategy set S_i, is represented by a vertex of a weighted graph $G = (V, E)$ (called *social graph*). To each edge $e = (i, j) \in E$, whose weight is denoted by w_e, is linked a two-players game $\mathcal{G}_e$ with potential function Φ_e in which the set of strategies of endpoints are exactly S_i and S_j. We denote with u_i^e the utility function of player i in the game $\mathcal{G}_e$. Given a strategy profile $\mathbf{x}$, the utility function of player i in the local interaction game $\mathcal{G}$ sets

$$u_i(\mathbf{x}) = \sum_{e=(i,j)} w_e \cdot u_i^e(x_i, x_j)\,.$$

It is easy to check that the function $\Phi = \sum_e w_e {\cdot} \Phi_e$ is a potential function for the local interaction game $\mathcal{G}$.

Logit choice function. We study the interaction of n players of a strategic game $\mathcal{G}$ that update their strategy according to the *logit choice function* [28, 9, 33] described as follows: from profile $\mathbf{x} \in S$ player $i \in [n]$ updates her strategy to $y \in S_i$ with probability

$$\sigma_i(y \mid \mathbf{x}) = \frac{e^{\beta u_i(\mathbf{x}_{-i}, y)}}{\sum_{z \in S_i} e^{\beta u_i(\mathbf{x}_{-i}, z)}}\,. \tag{1}$$

In other words, the logit choice function leans towards strategies promising higher utility. The parameter $\beta \geqslant 0$ is a measure of how much the utility influences the choice of the player.

All-logit. In this paper we consider the *all-logit* dynamics, by which *all* players *concurrently* update their strategy using the logit choice function. Most of the previous works have focused on dynamics where at each step *one* player is chosen uniformly at random and she updates her strategy by following the logit choice function. We call that dynamics *one-logit*, to distinguish it from the *all-logit.*

The all-logit dynamics induces a Markov chain over the set of strategy profiles whose transition probability $P(\mathbf{x}, \mathbf{y})$ from profile $\mathbf{x} = (x_1, \ldots, x_n)$ to profile $\mathbf{y} = (y_1, \ldots, y_n)$ is

$$P(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} \sigma_i(y_i \mid \mathbf{x}) = \frac{e^{\beta \sum_{i=1}^{n} u_i(\mathbf{x}_{-i}, y_i)}}{\prod_{i=1}^{n} \sum_{z \in S_i} e^{\beta u_i(\mathbf{x}_{-i}, z)}}\,. \tag{2}$$

Sometimes it is useful to write the transition probability from $\mathbf{x}$ to $\mathbf{y}$ in terms of the *cumulative utility* of $\mathbf{x}$ with respect to $\mathbf{y}$ defined as $U(\mathbf{x}, \mathbf{y}) = \sum_i u_i(\mathbf{x}_{-i}, y_i)$. Indeed, by observing that

$$\prod_{i=1}^{n} \sum_{z \in S_i} e^{\beta u_i(\mathbf{x}_{-i}, z)} = \sum_{\mathbf{z} \in S} \prod_{i=1}^{n} e^{\beta u_i(\mathbf{x}_{-i}, z_i)}\,,$$

we can rewrite (2) as

$$P(\mathbf{x}, \mathbf{y}) = \frac{e^{\beta U(\mathbf{x},\mathbf{y})}}{D(\mathbf{x})}\,, \tag{3}$$

where $D(\mathbf{x}) = \sum_{\mathbf{z} \in S} e^{\beta U(\mathbf{x},\mathbf{z})}$. For a potential game $\mathcal{G}$ with potential Φ, we can define the *cumulative potential* of $\mathbf{x}$ with respect to $\mathbf{y}$ as $\Psi(\mathbf{x}, \mathbf{y}) = \sum_i \Phi(\mathbf{x}_{-i}, y_i)$. Simple algebraic manipulations show that, for a potential game, we can rewrite the transition probabilities in (3) as

$$P(\mathbf{x}, \mathbf{y}) = \frac{e^{-\beta \Psi(\mathbf{x},\mathbf{y})}}{T(\mathbf{x})}\,,$$

where $T(\mathbf{x}) = \sum_{\mathbf{z} \in S} e^{-\beta \Psi(\mathbf{x},\mathbf{z})}$.

It is easy to see that a Markov chain with transition matrix (2) is ergodic. Indeed, for example, ergodicity follows from the fact that all entries of the transition matrix are strictly positive.

Reversibility, Observables, Mixing time. In this work we focus on three features of the all-logit dynamics, that we formally define here.

Let $\mathcal{M}$ be a Markov chain with transition matrix P and state set S. $\mathcal{M}$ is *reversible* with respect to a distribution π if, for every pair of states $x, y \in S$, the following *detailed balance condition* holds

$$\pi(x)P(x, y) = \pi(y)P(y, x)\,. \tag{4}$$

It is easy to see that if $\mathcal{M}$ is reversible with respect to π then π is also stationary.

An *observable* O is a function $O \colon S \to \mathbb{R}$, i.e. it is a function that assigns a value to each strategy profile of the game.

An ergodic Markov chain has a unique stationary distribution π and for every starting profile $\mathbf{x}$ the distribution $P^t(\mathbf{x}, \cdot)$ of the chain at time t converges to π as t goes to infinity. The *mixing time* is a measure of how long it takes to get close to the stationary distribution from the *worst-case* starting profile

$$t_{\mathrm{mix}}(\varepsilon) = \inf\left\{ t \in \mathbb{N} \,:\, \|P^t(\mathbf{x}, \cdot) - \pi\|_{\mathrm{TV}} \leqslant \varepsilon \text{ for all } \mathbf{x} \in S \right\}\,,$$

where $\|P^t(\mathbf{x}, \cdot) - \pi\|_{\mathrm{TV}} = \frac{1}{2} \sum_{\mathbf{y} \in S} |P^t(\mathbf{x}, \mathbf{y}) - \pi(\mathbf{y})|$ is the *total variation distance*. We will usually use t_{mix} for $t_{\mathrm{mix}}(1/4)$. We refer the reader to [26] for a more detailed description of notational conventions about Markov chains and mixing times.

3 Warm-up: two-player games

In this section we compare the behavior of the one- and the all-logit dynamics for two simple two-player potential games (thus two simple local information games): a *coordination game* and the *Prisoner's Dilemma*. The analysis of these games highlights that the stationary distribution of the two dynamics can significantly differ. However, it turns out that for both games the Markov chain induced by the all-logit is reversible, just as for the one-logit dynamics. More surprisingly, we see that the expected number of players taking a certain action in each one of these games is exactly the same regardless whether the expectation is taken according the stationary distribution of the all-logit or of the one-logit. Finally, we observe that the mixing time of the all-logit dynamics is asymptotically the same than the mixing time of the one-logit. Next sections will show that these results are not accidental.

Two-player coordination games. These are games in which the players have an advantage in selecting the same strategy. They are often used to model the spread of a new technology [35]: two players have to decide whether to adopt or not a new technology. Each player prefers to adopt the same technology as the other player. We denote by 0 the strategy of adopting the new technology and by 1 the strategy of adopting the old technology. The game is formally described by the following payoff matrix

$$\begin{array}{c|c|c|} & 0 & 1 \\ \hline 0 & a,a & c,d \\ \hline 1 & d,c & b,b \\ \hline \end{array} \tag{5}$$

We assume that $a > d$ and $b > c$ (meaning that players prefer to coordinate) and that $a - d = b - c = \Delta$ (meaning that there is not a risk dominant strategy [17]). It is easy to see that this game is a potential game. It is well known that the stationary distribution of the one-logit of a potential game is the Gibbs distribution, that assigns to $\mathbf{x} \in S$ probability $e^{-\beta\Phi(\mathbf{x})}/Z$, where $Z = \sum_{\mathbf{x}\in S} e^{-\beta\Phi(\mathbf{x})}$ is the *partition function*.

The transition matrix of the Markov chain induced by the all-logit dynamics is

$$P = \left(\begin{array}{c|cccc} & 00 & 01 & 10 & 11 \\ \hline 00 & (1-p)^2 & p(1-p) & p(1-p) & p^2 \\ 01 & (1-p)p & p^2 & (1-p)^2 & (1-p)p \\ 10 & p(1-p) & (1-p)^2 & p^2 & p(1-p) \\ 11 & p^2 & p(1-p) & p(1-p) & (1-p)^2 \end{array}\right)$$

where $p = 1/(1 + e^{\Delta\beta})$. Observe that this transition matrix is doubly-stochastic, that implies that the stationary distribution of the all-logit is uniform (and hence very different from the one-logit case). However, it is easy to check that the chain is reversible and the mixing time is $\Theta\left(e^{\Delta\beta}\right)$ (as in the one-logit case). Moreover, the expected number of players adopting the new strategy at stationarity is 1, both when considering the one- and the all-logit dynamics.

Prisoner's Dilemma. The Prisoner's Dilemma game is described by the payoff matrix given in (5), where with 0 we denote the strategy `Confess` and with 1 the strategy `Defect`. Moreover, payoffs satisfy the following conditions: (i) $a > d$ (so that 00 is a Nash equilibrium); (ii) $b < c$ (so that 11 is not a Nash equilibrium); (iii) $2a < c + d < 2b$ (so that 11 is the social optimum and 00 is the worst social profile). It is easy to check that the game is a potential game.

The transition matrix of the Markov chain induced by the all-logit dynamics is

$$P = \left(\begin{array}{c|cccc} & 00 & 01 & 10 & 11 \\ \hline 00 & (1-p)^2 & p(1-p) & p(1-p) & p^2 \\ 01 & (1-p)(1-q) & p(1-q) & q(1-p) & pq \\ 10 & (1-p)(1-q) & q(1-p) & p(1-q) & pq \\ 11 & (1-q)^2 & q(1-q) & q(1-q) & q^2 \end{array}\right)$$

where we let $p = 1/(1 + e^{(a-d)\beta})$ be the probability a player does not confess given the other player is currently confessing and $q = 1/(1+e^{(c-b)\beta})$ be the probability a player does not confess given the other player is currently not confessing. Note that both p and q go to 0 as β goes to infinity.

It is easy to check that the transition matrix is reversible (as for the one-logit). The stationary distribution is

$$\pi(CC) = \frac{(1-q)^2}{(1+p-q)^2} \qquad \pi(NN) = \frac{p^2}{(1+p-q)^2} \qquad \pi(NC) = \pi(CN) = \frac{p(1-q)}{(1+p-q)^2}\,.$$

Moreover, we can see that that the mixing time is upper bounded by a constant independent of β (as for the one-logit). You may also check that the expected number of confessing prisoners is exactly the same in the stationary distribution of the one- and of the all-logit.

4 Reversibility and stationary distribution

Reversibility is an important property of Markov chains and, in general, of stochastic processes. Roughly speaking, for a reversible Markov chain the stationary frequency of transitions from a state x to a state y is equal to the stationary frequency of transitions from y to x. It is easy to see that the one-logit for a game $\mathcal{G}$ is reversible if and only if $\mathcal{G}$ is a potential game. This does not hold for the all-logit. Indeed, we will prove that the class of games for which the all-logit is reversible is exactly the class of local interaction games.

4.1 Reversibility criteria

As previously stated, a Markov chain $\mathcal{M}$ is reversible if it satisfies the detailed balance condition (4). The Kolmogorov reversibility criterion allows us to establish the reversibility of a process directly from the transition probabilities. Before stating the criterion, we introduce the following notation. A *directed path* Γ from state $x \in S$ to state $y \in S$ is a sequence of states $\langle x_0, x_1, \ldots, x_\ell \rangle$ such that $x_0 = x$ and $x_\ell = y$. The probability $\mathbf{P}(\Gamma)$ of path Γ is defined as $\mathbf{P}(\Gamma) = \prod_{j=1}^{\ell} P(x_{j-1}, x_j)$. The *inverse of path* $\Gamma = \langle x_0, x_1, \ldots, x_\ell \rangle$ is the path $\Gamma^{-1} = \langle x_\ell, x_{\ell-1}, \ldots, x_0 \rangle$. Finally, a cycle C is simply a path from a state x to itself. We are now ready to state Kolmogorov's reversibility criterion (see, for example, [22]).

Theorem 1 (Kolmogorov's Reversibility Criterion)**.** *An irreducible Markov chain $\mathcal{M}$ with state space S and transition matrix P is reversible if and only if for every cycle C it holds that*

$$\mathbf{P}(C) = \mathbf{P}\left(C^{-1}\right).$$

The following lemma will be very useful for proving reversibility conditions for the all-logit dynamics and for stating a closed expression for its stationary distribution.

Lemma 2. *Let $\mathcal{M}$ be an irreducible Markov chain with transition probability P and state space S. $\mathcal{M}$ is reversible if and only if for every pair of states $x, y \in S$, there exists a constant $c_{x,y}$ such that for all paths Γ from x to y, it holds that*

$$\frac{\mathbf{P}(\Gamma)}{\mathbf{P}(\Gamma^{-1})} = c_{x,y}\,.$$

Proof. Fix $x, y \in S$ and consider two paths, Γ_1 and Γ_2, from x to y. Let C_1 and C_2 be the cycles $C_1 = \Gamma_1 \circ \Gamma_2^{-1}$ and $C_2 = \Gamma_2 \circ \Gamma_1^{-1}$, where $\circ$ denotes the concatenation of paths. If $\mathcal{M}$ is reversible then, by the Kolmogorov Reversibility Criterion, $\mathbf{P}(C_1) = \mathbf{P}(C_2)$. On the other hand,

$$\mathbf{P}(C_1) = \mathbf{P}(\Gamma_1) \cdot \mathbf{P}\left(\Gamma_2^{-1}\right) \quad \text{and} \quad \mathbf{P}(C_2) = \mathbf{P}(\Gamma_2) \cdot \mathbf{P}\left(\Gamma_1^{-1}\right).$$

Thus

$$\frac{\mathbf{P}(\Gamma_1)}{\mathbf{P}\left(\Gamma_1^{-1}\right)} = \frac{\mathbf{P}(\Gamma_2)}{\mathbf{P}\left(\Gamma_2^{-1}\right)}\,.$$

For the other direction, fix $z \in S$ and, for all $x \in S$, set $\tilde{\pi}(x) = c_{z,x}/Z$, where $Z = \sum_x c_{z,x}$ is the normalizing constant. Now consider any two states $x, y \in S$ of $\mathcal{M}$, let Γ_1 be any path from z to x and and set $\Gamma_2 = \Gamma_1 \circ \langle x, y \rangle$ (that is, Γ_2 is Γ_1 concatenated with the edge (x, y)). We have that

$$\begin{aligned}
\frac{\tilde{\pi}(x)}{\tilde{\pi}(y)} &= \frac{c_{z,x}}{c_{z,y}} \\
&= \frac{\mathbf{P}(\Gamma_1)}{\mathbf{P}\left(\Gamma_1^{-1}\right)} \cdot \frac{\mathbf{P}(\Gamma_2)}{\mathbf{P}\left(\Gamma_2^{-1}\right)} \\
&= \frac{\mathbf{P}(\Gamma_1)}{\mathbf{P}\left(\Gamma_1^{-1}\right)} \cdot \frac{\mathbf{P}\left(\Gamma_1^{-1}\right) \cdot P(y,x)}{\mathbf{P}(\Gamma_1) \cdot P(x,y)} \\
&= \frac{P(y,x)}{P(x,y)}
\end{aligned}$$

and therefore $\mathcal{M}$ is reversible with respect to $\tilde{\pi}$. □

4.2 All-logit reversibility implies potential games

In this section we prove that if the all-logit for a game $\mathcal{G}$ is reversible then $\mathcal{G}$ is a potential game.

The following lemma shows a condition on the cumulative utility of a game $\mathcal{G}$ that is necessary and sufficient for the reversibility of the all-logit of $\mathcal{G}$.

Lemma 3. *The all-logit for game $\mathcal{G}$ is reversible if and only if the following property holds for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in S$:*

$$U(\mathbf{x},\mathbf{y}) - U(\mathbf{y},\mathbf{x}) = \Big(U(\mathbf{x},\mathbf{z}) + U(\mathbf{z},\mathbf{y})\Big) - \Big(U(\mathbf{y},\mathbf{z}) + U(\mathbf{z},\mathbf{x})\Big). \tag{6}$$

Proof. To prove the only if part, pick any three $\mathbf{x}, \mathbf{y}, \mathbf{z} \in S$ and consider paths $\Gamma_1 = \langle \mathbf{x}, \mathbf{y} \rangle$ $\Gamma_2 = \langle \mathbf{x}, \mathbf{z}, \mathbf{y} \rangle$. From Lemma 2 we have that reversibility implies

$$\frac{\mathbf{P}(\Gamma_1)}{\mathbf{P}\left(\Gamma_1^{-1}\right)} = \frac{\mathbf{P}(\Gamma_2)}{\mathbf{P}\left(\Gamma_2^{-1}\right)}$$

whence

$$\frac{e^{\beta U(\mathbf{x},\mathbf{y})}}{D(\mathbf{x})} \frac{D(\mathbf{y})}{e^{\beta U(\mathbf{y},\mathbf{x})}} = \frac{e^{\beta U(\mathbf{x},\mathbf{z})}}{D(\mathbf{x})} \frac{e^{\beta U(\mathbf{z},\mathbf{y})}}{D(\mathbf{z})} \frac{D(\mathbf{y})}{e^{\beta U(\mathbf{y},\mathbf{z})}} \frac{D(\mathbf{z})}{e^{\beta U(\mathbf{z},\mathbf{x})}}.$$

which in turn implies (6).

As for the if part, let us fix state $\mathbf{z} \in S$ and define $\tilde{\pi}(\mathbf{x}) = \frac{P(\mathbf{z},\mathbf{x})}{Z \cdot P(\mathbf{x},\mathbf{z})}$, where Z is the normalizing constant. For any $\mathbf{x}, \mathbf{y} \in S$, we have

$$\frac{\tilde{\pi}(\mathbf{x})}{\tilde{\pi}(\mathbf{y})} = \frac{P(\mathbf{z},\mathbf{x})}{P(\mathbf{x},\mathbf{z})} \cdot \frac{P(\mathbf{y},\mathbf{z})}{P(\mathbf{z},\mathbf{y})} = \frac{e^{\beta U(\mathbf{z},\mathbf{x})}}{e^{\beta U(\mathbf{x},\mathbf{z})}} \cdot \frac{e^{\beta U(\mathbf{y},\mathbf{z})}}{e^{\beta U(\mathbf{z},\mathbf{y})}} \cdot \frac{D(\mathbf{x})}{D(\mathbf{y})} = \frac{e^{\beta U(\mathbf{y},\mathbf{x})}}{e^{\beta U(\mathbf{x},\mathbf{y})}} \cdot \frac{D(\mathbf{x})}{D(\mathbf{y})} = \frac{P(\mathbf{y},\mathbf{x})}{P(\mathbf{x},\mathbf{y})},$$

where the first equality follows from the definition of $\tilde{\pi}$, the second and the fourth follow from (3) and the third follows from (6). Therefore, the detailed balance equation holds for $\tilde{\pi}$ and thus the Markov chain is reversible. □

We are now ready to prove that the all-logit is reversible only for potential games.

Theorem 4. *If the all-logit for game $\mathcal{G}$ is reversible then $\mathcal{G}$ is a potential game.*

Proof. We show that if the all-logit is reversible then the utility improvement $I(\Gamma)$ over any circuit Γ of length 4 is 0. The theorem then follows by Theorem 32 in Appendix A.

Consider circuit $\Gamma = \langle \mathbf{x}, \mathbf{z}, \mathbf{y}, \mathbf{w} \rangle$ and let i be the player in which $\mathbf{x}$ and $\mathbf{z}$ differ and let j be the player in which $\mathbf{z}$ and $\mathbf{y}$ differ. Then $\mathbf{y}$ and $\mathbf{w}$ differ in player i and $\mathbf{w}$ and $\mathbf{x}$ differ in player j. In other words, $\mathbf{z} = (\mathbf{x}_{-i}, y_i) = (\mathbf{y}_{-j}, x_j)$ and $\mathbf{w} = (\mathbf{x}_{-i}, y_j) = (\mathbf{y}_{-i}, x_i)$. Therefore we have that

$$
\begin{aligned}
U(\mathbf{x},\mathbf{y}) &= \textstyle\sum_{k \neq i,j} u_k(\mathbf{x}) + u_i(\mathbf{z}) + u_j(\mathbf{w}) & U(\mathbf{y},\mathbf{x}) &= \textstyle\sum_{k \neq i,j} u_k(\mathbf{y}) + u_i(\mathbf{w}) + u_j(\mathbf{z}) \\
U(\mathbf{x},\mathbf{z}) &= \textstyle\sum_{k \neq i,j} u_k(\mathbf{x}) + u_i(\mathbf{z}) + u_j(\mathbf{x}) & U(\mathbf{z},\mathbf{y}) &= \textstyle\sum_{k \neq i,j} u_k(\mathbf{z}) + u_i(\mathbf{z}) + u_j(\mathbf{y}) \\
U(\mathbf{y},\mathbf{z}) &= \textstyle\sum_{k \neq i,j} u_k(\mathbf{y}) + u_i(\mathbf{y}) + u_j(\mathbf{z}) & U(\mathbf{z},\mathbf{x}) &= \textstyle\sum_{k \neq i,j} u_k(\mathbf{z}) + u_i(\mathbf{x}) + u_j(\mathbf{z})
\end{aligned}
$$

By plugging the above expressions into (6) and rearranging terms, we obtain

$$
\Big(u_i(\mathbf{z}) - u_i(\mathbf{x})\Big) + \Big(u_j(\mathbf{y}) - u_j(\mathbf{z})\Big) + \Big(u_i(\mathbf{w}) - u_i(\mathbf{y})\Big) + \Big(u_j(\mathbf{x}) - u_j(\mathbf{w})\Big) = 0
$$

which shows $I(\Gamma) = 0$. □

4.3 A necessary and sufficient condition for all-logit reversibility

In the previous section we have established that the all-logit is reversible only for potential games and therefore, from now on, we only consider potential games $\mathcal{G}$ with potential function Φ. In this section we present in Theorem 6 a necessary and sufficient condition for reversibility that involves the potential and the cumulative potential. The condition will then be used in the next section to prove that local interaction games are exactly the games whose all-logit is reversible.

We start by re-writing Lemma 3 in terms of cumulative potential.

Lemma 5. *The all-logit is reversible if and only if for every* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in S$*:*

$$
\Psi(\mathbf{x},\mathbf{y}) - \Psi(\mathbf{y},\mathbf{x}) = \Big(\Psi(\mathbf{x},\mathbf{z}) + \Psi(\mathbf{z},\mathbf{y})\Big) - \Big(\Psi(\mathbf{y},\mathbf{z}) + \Psi(\mathbf{z},\mathbf{x})\Big). \tag{7}
$$

We are now ready to prove a necessary and sufficient condition for reversibility that involves potential and cumulative potential.

Theorem 6. *The all-logit for a game* $\mathcal{G}$ *with potential* Φ *is reversible if and only if, for all strategy profiles* $\mathbf{x}, \mathbf{y} \in S$*,*

$$
\Psi(\mathbf{x},\mathbf{y}) - \Psi(\mathbf{y},\mathbf{x}) = (n-2)\left(\Phi(\mathbf{x}) - \Phi(\mathbf{y})\right). \tag{8}
$$

Proof. Clearly (8) implies (7). As for the other direction, we proceed by induction on the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$. Let $\mathbf{x}$ and $\mathbf{y}$ be two profiles at Hamming distance 1; that is, $\mathbf{x}$ and $\mathbf{y}$ differ in only one player, say j. This implies that $(y_j, \mathbf{x}_{-j}) = \mathbf{y}$ and $(x_j, \mathbf{y}_{-j}) = \mathbf{x}$. Moreover, for $i \neq j$, $(y_i, \mathbf{x}_{-i}) = \mathbf{x}$ and $(x_i, \mathbf{y}_{-i}) = \mathbf{y}$. Thus,

$$
\begin{aligned}
\Psi(\mathbf{x},\mathbf{y}) - \Psi(\mathbf{y},\mathbf{x}) &= \sum_i \Big(\Phi(y_i, \mathbf{x}_{-i}) - \Phi(x_i, \mathbf{y}_{-i})\Big) \\
&= \Big(\Phi(y_j, \mathbf{x}_{-j}) - \Phi(x_j, \mathbf{y}_{-j})\Big) + \sum_{i \neq j} \Big(\Phi(y_i, \mathbf{x}_{-i}) - \Phi(x_i, \mathbf{y}_{-i})\Big) \\
&= \Big(\Phi(\mathbf{y}) - \Phi(\mathbf{x})\Big) + (n-1)\Big(\Phi(\mathbf{x}) - \Phi(\mathbf{y})\Big) = (n-2)\Big(\Phi(\mathbf{x}) - \Phi(\mathbf{y})\Big).
\end{aligned}
$$

Now assume that the claim holds for any pair of profiles at Hamming distance $k < n$ and let $\mathbf{x}$ and $\mathbf{y}$ be two profiles at distance $k+1$. Let j be any player such that $x_j \neq y_j$ and let

$\mathbf{z} = (y_j, \mathbf{x}_{-j})$: $\mathbf{z}$ is at distance at most k from $\mathbf{x}$ and from $\mathbf{y}$. Then, by (7) and by the inductive hypothesis, we have

$$\begin{aligned}\Psi(\mathbf{x},\mathbf{y}) - \Psi(\mathbf{y},\mathbf{x}) &= \Big(\Psi(\mathbf{x},\mathbf{z}) + \Psi(\mathbf{z},\mathbf{y})\Big) - \Big(\Psi(\mathbf{y},\mathbf{z}) + \Psi(\mathbf{z},\mathbf{x})\Big) \\ &= (n-2)\Big(\Phi(\mathbf{x}) + \Phi(\mathbf{z}) - \Phi(\mathbf{y}) - \Phi(\mathbf{z})\Big) = (n-2)\Big(\Phi(\mathbf{x}) - \Phi(\mathbf{y})\Big)\,. \qquad \square\end{aligned}$$

4.4 Reversibility and local interaction games

Here we prove that the games whose all-logit is reversible are exactly the local interaction games.

A potential $\Phi : S_1 \times \cdots \times S_n \to \mathbb{R}$ is a *two-player potential* if there exist $u, v \in [n]$ such that, for any $\mathbf{x}, \mathbf{y} \in S$ with $x_u = y_u$ and $x_v = y_v$ we have $\Phi(\mathbf{x}) = \Phi(\mathbf{y})$. In other words, Φ is a function of only its u-th and v-th argument. An interesting fact about two-player potential games is given by the following lemma.

Lemma 7. *Any two-player potential satisfies* (8)*.*

Proof. Let Φ be a two-player potential and let u and v be its two players. Then we have that for $w \neq u, v$, $\Phi(y_w, \mathbf{x}_{-w}) = \Phi(\mathbf{x})$ and that $\Phi(y_u, \mathbf{x}_{-u}) = \Phi(x_v, \mathbf{y}_{-v})$ and $\Phi(y_v, \mathbf{x}_{-v}) = \Phi(x_u, \mathbf{y}_{-u})$. Thus

$$\Psi(\mathbf{x},\mathbf{y}) = \Phi(y_u, \mathbf{x}_{-u}) + \Phi(y_v, \mathbf{x}_{-v}) + (n-2)\Phi(\mathbf{x})$$

and

$$\begin{aligned}\Psi(\mathbf{y},\mathbf{x}) &= \Phi(x_v, \mathbf{y}_{-v}) + \Phi(x_u, \mathbf{y}_{-u}) + (n-2)\Phi(\mathbf{y}) \\ &= \Phi(y_u, \mathbf{x}_{-u}) + \Phi(y_v, \mathbf{x}_{-v}) + (n-2)\Phi(\mathbf{y})\,. \qquad \square\end{aligned}$$

We say that a potential Φ is the sum of two-player potentials if there exist N two-player potentials $\Phi_1, \ldots, \Phi_N$ such that $\Phi = \Phi_1 + \cdots + \Phi_N$. It is easy to see that generality is not lost by further requiring that $1 \leqslant l \neq l' \leqslant N$ implies $(u_l, v_l) \neq (u_{l'}, v_{l'})$, where u_l and v_l are the two players of potential Φ_l. At every game $\mathcal{G}$ whose potential is the sum of two-player potentials, i.e., $\Phi = \Phi_1 + \cdots + \Phi_N$, we can associate a *social graph* G that has a vertex for each player of $\mathcal{G}$ and has edge (u, v) iff there exists l such that potential Φ_l depends on players u and v. In other words, each game whose potential is the sum of two-player potentials is a local interaction game[1].

Observe that the sum of two potentials satisfying (8) also satisfies (8). Hence we have the following theorem.

Theorem 8. *The all-logit dynamics for a local interaction game is reversible.*

Next we prove that if an n-player potential Φ satisfies (8) then it can be written as the sum of at most $N = \binom{n}{2}$ two-player potentials, $\Phi_1, \ldots, \Phi_N$ and thus it represents a local interaction game. We do so by describing an effective procedure that constructs the N two-player potentials.

Without loss of generality, we assume that each strategy set S_i includes strategy 0 and denote by $\mathbf{0}$ the strategy profile consisting of n 0's. Moreover, we fix an arbitrary ordering $(u_1, v_1), \ldots, (u_N, v_N)$ of the N unordered pairs of players. For a potential Φ we define the sequence $\vartheta_0, \ldots, \vartheta_N$ of potentials as follows: $\vartheta_0 = \Phi$ and, for $i = 1, \ldots, N$, set

$$\vartheta_i = \vartheta_{i-1} - \Phi_i \tag{9}$$

[1]Note that we can assume without loss of generality that all edge weights are 1 as we can scale down the weights and scale up potential function values without changing the behavior of the dynamics.

where, for $\mathbf{x} \in S$, $\Phi_i(\mathbf{x})$ is defined as

$$\Phi_i(\mathbf{x}) = \vartheta_{i-1}(x_{u_i}, x_{v_i}, \mathbf{0}_{-u_i v_i}) \,.$$

Observe that, for $i = 1, \ldots, N$, Φ_i is a two-player potential and its players are u_i and v_i. From Lemma 7, Φ_i satisfies (8). Hence, if Φ satisfies (8), then also ϑ_i, for $i = 1, \ldots, N$, satisfies (8).

By summing for $i = 1, \ldots, N$ in (9) we obtain

$$\sum_{i=1}^{N} \vartheta_i = \sum_{i=0}^{N-1} \vartheta_i - \sum_{i=1}^{N} \Phi_i \,.$$

Thus

$$\Phi - \vartheta_N = \sum_{i=1}^{N} \Phi_i \,.$$

The next two lemmas prove that, if Φ satisfies (8), then ϑ_N is identically zero. This implies that Φ is the sum of at most N non-zero two-player potentials and thus a local interaction game.

A *ball* $B(r, \mathbf{x})$ of radius $r \leqslant n$ centered in $\mathbf{x} \in S$ is the subset of S containing all profiles $\mathbf{y}$ that differ from $\mathbf{x}$ in at most r coordinates.

Lemma 9. *For any n-player potential function Φ and for any ordering of the pairs of players, $\vartheta_N(\mathbf{x}) = 0$ for every $\mathbf{x} \in B(2, \mathbf{0})$.*

Proof. We distinguish three cases based on the distance of $\mathbf{x}$ from $\mathbf{0}$.
$\underline{\mathbf{x} = \mathbf{0}}$: for every $i \geqslant 1$, we have

$$\vartheta_i(\mathbf{0}) = \vartheta_{i-1}(\mathbf{0}) - \Phi_i(\mathbf{0}) = \vartheta_{i-1}(\mathbf{0}) - \vartheta_{i-1}(\mathbf{0}) = 0 \,.$$

<u>$\mathbf{x}$ is at distance 1 from 0:</u> That is, there exists $u \in [n]$ such that $\mathbf{x} = (x_u, \mathbf{0}_{-u})$, with $x_u \neq 0$. Let us denote by $t(u)$ the smallest t such that the t-th pair contains u. We next show that for $i \geqslant t(u)$, $\vartheta_i(\mathbf{x}) = 0$. Indeed, we have that if u is a component of the i-th pair then

$$\vartheta_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \Phi_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \vartheta_{i-1}(\mathbf{x}) = 0 \,;$$

On the other hand, if u is not a component of the i-th pair then

$$\vartheta_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \Phi_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \vartheta_{i-1}(\mathbf{0}) = \vartheta_{i-1}(\mathbf{x}) \,;$$

<u>$\mathbf{x}$ is at distance 2 from 0:</u> That is, there exist u and v such that $\mathbf{x} = (x_u, x_v, \mathbf{0}_{-uv})$, with $x_u, x_v \neq 0$.
Let t be the index of the pair (u, v). Notice that $t \geqslant t(u), t(v)$. We show that $\vartheta_t(\mathbf{x}) = 0$ and that this value does not change for all $i > t$. Indeed, we have

$$\vartheta_t(\mathbf{x}) = \vartheta_{t-1}(\mathbf{x}) - \Phi_t(\mathbf{x}) = \vartheta_{t-1}(\mathbf{x}) - \vartheta_{t-1}(\mathbf{x}) = 0 \,;$$

If instead neither of u and v belongs to the i-th pair, with $i > t$, then we have

$$\vartheta_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \Phi_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \vartheta_{i-1}(\mathbf{0}) = \vartheta_{i-1}(\mathbf{x}) \,;$$

Finally, suppose that the i-th pair, for $i > t$, contains exactly one of u and v, say u. Then we have

$$\vartheta_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \Phi_i(\mathbf{x}) = \vartheta_{i-1}(\mathbf{x}) - \vartheta_{i-1}(x_u, \mathbf{0}_{-u}).$$

We conclude the proof by observing that $i - 1 \geq t \geqslant t(u)$ and thus, by the previous case, $\vartheta_{i-1}(x_u, \mathbf{0}_{-u}) = 0$. □

The next lemma shows that if a potential ϑ satisfies (8) and is constant in a ball of radius 2, then it is constant everywhere.

Lemma 10. *Let ϑ be a function that satisfies* (8)*. If there exist $\mathbf{x} \in S$ and $c \in \mathbb{R}$ such that $\vartheta(\mathbf{y}) = c$ for every $\mathbf{y} \in B(2, \mathbf{x})$, then $\vartheta(\mathbf{y}) = c$ for every $\mathbf{y} \in S$.*

Proof. Fix $h > 2$ and suppose that $\vartheta(\mathbf{z}) = c$ for every $\mathbf{z} \in B(h-1, \mathbf{x})$. Consider $\mathbf{y} \in B(h, \mathbf{x}) \setminus B(h-1, \mathbf{x})$ and observe that $(y_i, \mathbf{x}_{-i}) \in B(h-1, \mathbf{x})$ and $(x_i, \mathbf{y}_{-i}) \in B(h-1, \mathbf{x})$ for every i such that $x_i \neq y_i$. It is easy to see that

$$(h-2)\left(\vartheta(\mathbf{x}) - \vartheta(\mathbf{y})\right) = \sum_{i\colon x_i \neq y_i} \Big(\vartheta(y_i, \mathbf{x}_{-i}) - \vartheta(x_i, \mathbf{y}_{-i}) \Big) = 0\,,$$

that implies $\vartheta(\mathbf{y}) = \vartheta(\mathbf{x}) = c$. □

We can thus conclude that if the all-logit of a potential game $\mathcal{G}$ is reversible then $\mathcal{G}$ is a local interaction game. By combining this result with Theorem 4 and Theorem 8, we obtain

Theorem 11. *The all-logit of game $\mathcal{G}$ is reversible if and only if $\mathcal{G}$ is a local interaction game.*

4.5 Stationary distribution for the all-logit of local interaction games

Theorem 12 (Stationary distribution)**.** *Let $\mathcal{G}$ be a local interaction game with potential function Φ. Then the stationary distribution of the all-logit for $\mathcal{G}$ is*

$$\pi(\mathbf{x}) \propto e^{(n-2)\beta\Phi(\mathbf{x})} \cdot T(\mathbf{x}) \tag{10}$$

where $T(\mathbf{x}) = \sum_{\mathbf{z} \in S} e^{-\beta\Psi(\mathbf{x},\mathbf{z})}$.

Proof. Fix any profile $\mathbf{y}$. The detailed balance equation gives for every $\mathbf{x} \in S$

$$\frac{\pi(\mathbf{x})}{\pi(\mathbf{y})} = \frac{P(\mathbf{y}, \mathbf{x})}{P(\mathbf{x}, \mathbf{y})} = e^{\beta(\Psi(\mathbf{x},\mathbf{y}) - \Psi(\mathbf{y},\mathbf{x}))} \frac{T(\mathbf{x})}{T(\mathbf{y})}\,.$$

By Theorem 6 we have

$$\pi(\mathbf{x}) = e^{(n-2)\beta\Phi(\mathbf{x})} \cdot T(\mathbf{x}) \left(\frac{\pi(\mathbf{y})}{e^{(n-2)\beta\Phi(\mathbf{y})} \cdot T(\mathbf{y})} \right)\,.$$

Since the term in parenthesis does not depend on $\mathbf{x}$ the theorem follows. □

Note that for a local interaction game $\mathcal{G}$ with potential function Φ. We write $\pi_1(\mathbf{x})$, the stationary distribution of the one-logit of $\mathcal{G}$, as $\pi_1(\mathbf{x}) = \gamma_1(\mathbf{x})/Z_1$ where $\gamma_1(\mathbf{x}) = e^{-\beta\Phi(\mathbf{x})}$ is the Boltzmann factor and $Z_1 = \sum_{\mathbf{x}} \gamma_1(\mathbf{x})$ is the partition function. From Theorem 12, we derive that $\pi_A(\mathbf{x})$, the stationary distribution of the all-logit of $\mathcal{G}$, can be written in similar way, i.e., $\pi_A(\mathbf{x}) = \frac{\gamma_A(\mathbf{x})}{Z_A}$, where

$$\gamma_A(\mathbf{x}) = \sum_{\mathbf{y} \in S} e^{-\beta[\Psi(\mathbf{x},\mathbf{y}) - (n-2)\Phi(\mathbf{x})]}$$

and $Z_A = \sum_{\mathbf{x} \in S} \gamma_A(\mathbf{x})$ is the partition function of the all-logit. Simple algebraic manipulation shows that, by setting

$$K(\mathbf{x}, \mathbf{y}) = 2 \cdot \Phi(\mathbf{x}) + \sum_{i \in [n]} \mathbf{d}_{\mathbf{x},\mathbf{y}}(i) \cdot \left(\Phi(\mathbf{x}_{-i}, y_i) - \Phi(\mathbf{x})\right)$$

where $\mathbf{d}_{\mathbf{x},\mathbf{y}}$ is the characteristic vector of positions i in which $\mathbf{x}$ and $\mathbf{y}$ differ (i.e., $\mathbf{d}_{\mathbf{x},\mathbf{y}}(i) = 1$ if $x_i \neq y_i$ and 0 otherwise), we can write $\gamma_A(\mathbf{x})$ and Z_A as

$$\gamma_A(\mathbf{x}) = \sum_{\mathbf{y}\in S} e^{-\beta K(\mathbf{x},\mathbf{y})} \qquad \text{and} \qquad Z_A = \sum_{\mathbf{x},\mathbf{y}} e^{-\beta K(\mathbf{x},\mathbf{y})}. \tag{11}$$

Furthermore, we can decompose $K(\mathbf{x}, \mathbf{y})$ in the contributions of each edge of the social graph G of $\mathcal{G}$. Specifically, we have the following lemma.

Lemma 13. *Let $\mathcal{G}$ be a local interaction game over social graph G. Then we have $K(\mathbf{x}, \mathbf{y}) = \sum_e K_e(\mathbf{x}, \mathbf{y})$, where for each edge $e = (u, v)$ of G, $K_e(\mathbf{x}, \mathbf{y})$ is defined as*

$$K_e(\mathbf{x}, \mathbf{y}) = 2\Phi_e(x_u, x_v) + \kappa_e(\mathbf{x}, \mathbf{y}) \tag{12}$$

and

$$\kappa_e(\mathbf{x}, \mathbf{y}) = \mathbf{d}(u)(\Phi_e(y_u, x_v) - \Phi_e(x_u, x_v)) + \mathbf{d}(v)(\Phi_e(x_u, y_v) - \Phi_e(x_u, x_v)).$$

Proof. Observe that $\sum_{e=(u,v)} 2 \cdot \Phi_e(x_u, x_v) = 2\Phi(\mathbf{x})$. Moreover we have

$$\sum_{e=(u,v)} \mathbf{d}(u)(\Phi_e(y_u, x_v) - \Phi_e(x_u, x_v)) = \sum_{i\in[n]} \mathbf{d}(i) \sum_{\substack{e=(u,v)\\ i=u}} (\Phi_e(y_u, x_v) - \Phi_e(x_u, x_v)).$$

Then we have that

$$\begin{aligned} \sum_e \kappa_e(\mathbf{x}, \mathbf{y}) = \sum_{i\in[n]} \mathbf{d}(i) \Bigg[&\sum_{\substack{e=(u,v)\\ i=u}} (\Phi_e(y_u, x_v) - \Phi_e(x_u, x_v)) + \sum_{\substack{e=(u,v)\\ i=v}} (\Phi_e(x_u, y_v) - \Phi_e(x_u, x_v)) \\ &+ \sum_{\substack{e=(u,v)\\ i\neq\{u,v\}}} (\Phi_e(x_u, x_v) - \Phi_e(x_u, x_v)) \Bigg] \\ = \sum_{i\in[n]} \mathbf{d}(i) \cdot (\Phi(\mathbf{x}_{-i}, y_i) - \Phi(\mathbf{x}))& . \end{aligned}$$

Hence,

$$\sum_{e=(u,v)} K_e(\mathbf{x}, \mathbf{y}) = \sum_{e=(u,v)} 2 \cdot \Phi_e(x_u, x_v) + \sum_e \kappa_e(\mathbf{x}, \mathbf{y}) = K(\mathbf{x}, \mathbf{y}).$$

□

5 Observables of local information games

In this section we study observables of local interaction games and we focus on the relation between the expected value, $\langle O, \pi_1 \rangle$, of an observable O at the stationarity of the one-logit and its expected value, $\langle O, \pi_A \rangle$, at the stationarity of the all-logit dynamics. We start by studying invariant observables; that is, observables for which the two expected values coincide. In Theorem 17, we give a sufficient condition for an observable to be invariant. The sufficient condition is related to the existence of a *decomposition* of the set $S \times S$ that decomposes the quantity K appearing in the expression for the stationary distribution of the all-logit of the local interaction game $\mathcal{G}$ (see Eq. 11) into a sum of two potentials. In Theorem 17 we show that if $\mathcal{G}$ admits such a decomposition μ and in addition observable O is also decomposed by μ (see Definition 16) then O has the same expected value at the stationarity of the one-logit and of the all-logit. We then go on to show that all local interaction games on *bipartite* social graphs admit a decomposition permutation (see Theorem 15) and give examples of invariant observables.

We then look at local interaction games $\mathcal{G}$ on general social graphs G and show that the expected values of a decomposable observable O with respect to the stationary distributions of the one-logit and of the all-logit differ by a quantity that depends on β and on how far away the social graph G is from being bipartite (which in turn is related to the smallest eigenvalue of G [32]).

The above findings follow from a relation between the partition functions of the one-logit and of the all-logit that might be of independent interest. More precisely, in Theorem 15 we show that if the game $\mathcal{G}$ admits a decomposition then the partition function of the all-logit is the square of the partition function of the one-logit. The partition function of the one-logit is easily seen to be equal to the partition function of the canonical ensemble used in Statistical Mechanics (see for example [24]). It is well known that a partition function of a canonical ensemble that is the union of two independent canonical ensembles is the product of the two partition functions. Thus Theorem 15 can be seen as a further confirmation that the all-logit can be decomposed into two independent one-logit dynamics.

Throughout this section we assume, for sake of ease of presentation, that each player has just two strategies available. Extending our results to any number of strategies is straightforward.

5.1 Decomposable observables for bipartite social graphs

We start by introducing the concept of a *decomposition* and then we define the concept of a *decomposable* observable.

Definition 14. *A permutation*

$$\mu\colon (\mathbf{x},\mathbf{y}) \mapsto (\mu_1(\mathbf{x},\mathbf{y}), \mu_2(\mathbf{x},\mathbf{y}))$$

of $S \times S$ is a decomposition *for a local interaction game $\mathcal{G}$ with potential Φ if, for all $(\mathbf{x},\mathbf{y})$, we have that*

$$K(\mathbf{x},\mathbf{y}) = \Phi(\mu_1(\mathbf{x},\mathbf{y})) + \Phi(\mu_2(\mathbf{x},\mathbf{y}))\,,$$

$\mu_1(\mathbf{x},\mathbf{y}) = \mu_2(\mathbf{y},\mathbf{x})$ and $\mu_2(\mathbf{x},\mathbf{y}) = \mu_1(\mathbf{y},\mathbf{x})$.

Theorem 15. *If a local interaction game $\mathcal{G}$ admits a decomposition μ then $Z_A = Z_1^2$.*

Proof. From (11) and from the fact that μ is a permutation of $S \times S$, we have

$$Z_A = \sum_{\mathbf{x},\mathbf{y}} e^{-\beta K(\mathbf{x},\mathbf{y})} = \sum_{\mathbf{x},\mathbf{y}} e^{-\beta[\Phi(\mu_1(\mathbf{x},\mathbf{y}))+\Phi(\mu_2(\mathbf{x},\mathbf{y}))]} = \sum_{\mathbf{x},\mathbf{y}} e^{-\beta[\Phi(\mathbf{x})+\Phi(\mathbf{y})]} = Z_1^2. \qquad \square$$

Definition 16. *An observable O is* decomposable *if there exists a decomposition μ such that, for all $(\mathbf{x},\mathbf{y})$, we have that*

$$O(\mathbf{x}) + O(\mathbf{y}) = O(\mu_1(\mathbf{x},\mathbf{y})) + O(\mu_2(\mathbf{x},\mathbf{y})).$$

We next prove that a decomposable observable has the same expectation at stationarity for the one-logit and the all-logit.

Theorem 17. *If observable O is decomposable then*

$$\langle O, \pi_1 \rangle = \langle O, \pi_A \rangle\,.$$

Proof. Suppose that O is decomposed by μ. Then we have that, for all $\mathbf{x} \in S$, $\gamma_A(\mathbf{x}) = \sum_{\mathbf{y}} \gamma_1(\mu_1(\mathbf{x},\mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x},\mathbf{y}))$ and thus

$$\begin{aligned}\langle O, \pi_A \rangle &= \frac{1}{Z_A} \sum_{\mathbf{x}} O(\mathbf{x}) \cdot \gamma_A(\mathbf{x}) \\ &= \frac{1}{Z_A} \sum_{\mathbf{x},\mathbf{y}} O(\mathbf{x}) \cdot \gamma_1(\mu_1(\mathbf{x},\mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x},\mathbf{y})) \\ &= \frac{1}{2} \cdot \frac{1}{Z_A} \sum_{\mathbf{x},\mathbf{y}} [O(\mathbf{x}) + O(\mathbf{y})] \cdot \gamma_1(\mu_1(\mathbf{x},\mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x},\mathbf{y}))\end{aligned}$$

where in the last equality we have used the fact that, $\mu_1(\mathbf{x},\mathbf{y}) = \mu_2(\mathbf{y},\mathbf{x})$ and $\mu_2(\mathbf{x},\mathbf{y}) = \mu_1(\mathbf{y},\mathbf{x})$ which implies that

$$\sum_{\mathbf{x},\mathbf{y}} O(\mathbf{x}) \cdot \gamma_1(\mu_1(\mathbf{x},\mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x},\mathbf{y})) = \sum_{\mathbf{x},\mathbf{y}} O(\mathbf{y}) \cdot \gamma_1(\mu_1(\mathbf{x},\mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x},\mathbf{y}))\,.$$

Now, since O is decomposable we have that $O(\mathbf{x}) + O(\mathbf{y}) = O(\mu_1(\mathbf{x},\mathbf{y})) + O(\mu_2(\mathbf{x},\mathbf{y}))$ and thus we can write

$$\begin{aligned}\langle O, \pi_A \rangle &= \frac{1}{2} \cdot \frac{1}{Z_A} \sum_{\mathbf{x},\mathbf{y}} [O(\mu_1(\mathbf{x},\mathbf{y})) + O(\mu_2(\mathbf{x},\mathbf{y}))] \cdot \gamma_1(\mu_1(\mathbf{x},\mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x},\mathbf{y})) \\ &= \frac{1}{2} \cdot \frac{1}{Z_A} \sum_{\mathbf{x},\mathbf{y}} [O(\mathbf{x}) + O(\mathbf{y})] \cdot \gamma_1(\mathbf{x}) \cdot \gamma_1(\mathbf{y}) \\ &= \frac{1}{Z_A} \sum_{\mathbf{x},\mathbf{y}} O(\mathbf{x}) \cdot \gamma_1(\mathbf{x}) \cdot \gamma_1(\mathbf{y}) \\ &= \sum_{\mathbf{x}} O(\mathbf{x}) \cdot \frac{\gamma_1(\mathbf{x})}{Z_1} \cdot \sum_{\mathbf{y}} \frac{\gamma_1(\mathbf{y})}{Z_1} \\ &= \sum_{\mathbf{x}} O(\mathbf{x}) \cdot \pi_1(\mathbf{x}) \cdot \sum_{\mathbf{y}} \pi_1(\mathbf{y}) \\ &= \langle O, \pi_1 \rangle\,.\end{aligned}$$

□

We next prove that for all local interaction games on a bipartite social graph there exists a decomposition. We start with the following sufficient condition for a permutation to be a decomposition.

Lemma 18. *Let $\mathcal{G}$ be a social interaction game on social graph G with potential Φ and let μ be a permutation of $S \times S$ such that, for all $\mathbf{x},\mathbf{y} \in S$, we have $\mu_1(\mathbf{x},\mathbf{y}) = \mu_2(\mathbf{y},\mathbf{x})$, $\mu_2(\mathbf{x},\mathbf{y}) = \mu_1(\mathbf{y},\mathbf{x})$ and for all edges $e = (u,v)$ of G and for all $\mathbf{x},\mathbf{y} \in S$ one of two equalities below holds.*

$$(\tilde{x}_u, \tilde{x}_v, \tilde{y}_u, \tilde{y}_v) = (x_u, y_v, y_u, x_v) \tag{13}$$

$$(\tilde{x}_u, \tilde{x}_v, \tilde{y}_u, \tilde{y}_v) = (y_u, x_v, x_u, y_v), \tag{14}$$

where $\tilde{\mathbf{x}} = \mu_1(\mathbf{x},\mathbf{y})$ and $\tilde{\mathbf{y}} = \mu_2(\mathbf{x},\mathbf{y})$. Then μ is a decomposition of $\mathcal{G}$.

Proof. We prove that $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(\tilde{x}_u, \tilde{x}_v) + \Phi_e(\tilde{y}_u, \tilde{y}_v)$ and the lemma is then obtained by summing over all edges e. We observe that, under both assignments described by (13) and (14), we have

$$\Phi_e(\tilde{x}_u, \tilde{x}_v) + \Phi_e(\tilde{y}_u, \tilde{y}_v) = \Phi_e(x_u, y_v) + \Phi_e(y_u, x_v)$$

and thus it is enough prove that $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(x_u, y_v) + \Phi_e(y_u, x_v)$.

We distinguish four cases:

1. $y_u = x_u$ and $y_v = x_v$. In this case, from Eq. 12, we have $K_e(\mathbf{x}, \mathbf{y}) = 2 \cdot \Phi_e(x_u, x_v)$ and thus $K_e(\mathbf{x}, \mathbf{y}) = \Phi_e(x_u, y_v) + \Phi_e(y_u, x_v)$.

2. $y_u \neq x_u$ and $y_v = x_v$. In this case, from Eq. 12, we have $K_e(\mathbf{x}, \mathbf{y}) = \Phi_e(x_u, x_v) + \Phi_e(y_u, x_v)$ and thus $K_e(\mathbf{x}, \mathbf{y}) = \Phi_e(x_u, y_v) + \Phi_e(y_u, x_v)$.

3. $y_u = x_u$ and $y_v \neq x_v$. In this case, from Eq. 12, we have $K_e(\mathbf{x}, \mathbf{y}) = \Phi_e(x_u, x_v) + \Phi_e(x_u, y_v)$ and thus $K_e(\mathbf{x}, \mathbf{y}) = \Phi_e(x_u, y_v) + \Phi_e(y_u, x_v)$.

4. $y_u \neq x_u$ and $y_v \neq x_v$. In this case, from Eq. 12, we have $K_e(\mathbf{x}, \mathbf{y}) = \Phi_e(y_u, x_v) + \Phi_e(x_u, y_v)$. □

We next prove that every social interaction game on a bipartite social graph admits a decomposition.

Theorem 19. *Let $\mathcal{G}$ be a social interaction game on a bipartite graph G. Then $\mathcal{G}$ admits a decomposition.*

Proof. Let (L, R) be the set of vertices in which G is bipartite. For each $(\mathbf{x}, \mathbf{y}) \in S \times S$ define $\tilde{\mathbf{x}} = \mu_1(\mathbf{x}, \mathbf{y})$ and $\tilde{\mathbf{y}} = \mu_2(\mathbf{x}, \mathbf{y})$ as follows: for every vertex u of G

- if $u \in L$, then set $\tilde{x}_u = x_u$ and $\tilde{y}_u = y_u$;
- if $u \in R$, then set $\tilde{x}_u = y_u$ and $\tilde{y}_u = x_u$.

First of all, observe that the mapping is an involution and thus it is also a permutation and that $\mu_1(\mathbf{x}, \mathbf{y}) = \mu_2(\mathbf{y}, \mathbf{x})$ and $\mu_2(\mathbf{x}, \mathbf{y}) = \mu_1(\mathbf{y}, \mathbf{x})$. Since G is bipartite, for every edge (u, v) exactly one endpoint is in L and exactly one is in R. If $u \in L$, then we have that $(\tilde{x}_u, \tilde{x}_v, \tilde{y}_u, \tilde{y}_v) = (x_u, y_v, y_u, x_v)$ and thus (13) is satisfied. If instead $u \in R$, then we have that $(\tilde{x}_u, \tilde{x}_v, \tilde{y}_u, \tilde{y}_v) = (y_u, x_v, x_u, y_v)$ and thus (14) is satisfied. Therefore for each edge one of Eq. 13 and 14 is satisfied. By Lemma 18, we can conclude that the mapping is a decomposition. □

We now give examples of decomposable observables.

The Diff observable. We start by looking at the observable Diff that returns the (signed) difference between the number of vertices adopting strategy 0 and the number of vertices adopting strategy 1. That is, $\mathsf{Diff}(\mathbf{x}) = n - 2\sum_u x_u$. In local interaction games used to model the diffusion of innovation in social networks and the spread of new technology (see, for example, [35]), this observable is a measure of how wide is the adoption of the innovation. The Diff observable is also meaningful in the Ising model for ferromagnetism (see, for example, [27]) as it is the measured magnetism.

To prove that Diff is decomposable we consider the mapping used in the proof of Theorem 19 and observe that, for every vertex u and for every $(\mathbf{x}, \mathbf{y}) \in S \times S$, we have $x_u + y_u = \tilde{x}_u + \tilde{y}_u$. Whence we conclude that $O(\mathbf{x}) + O(\mathbf{y}) = O(\tilde{\mathbf{x}}) + O(\tilde{\mathbf{y}})$.

The MonoC observable. Another interesting decomposable observable is the signed difference MonoC between the number of 0-monochromatic edges of the social graph (that is, edges in which both endpoints play 0) and the number of 1-monochromatic edges. That is, $\mathsf{MonoC}(\mathbf{x}) = \sum_{(u,v)\in E}(x_u + x_v - 1)$. Again, we consider the mapping of the proof of Theorem 19 and the decomposability of MonoC follows from the property that, for every $(\mathbf{x}, \mathbf{y}) \in S \times S$, we have $x_u + y_u = \tilde{x}_u + \tilde{y}_u$.

Theorem 20. *Observables* Diff *and* MonoC *are decomposable and thus, for local interaction games on bipartite social graphs,*

$$\langle \mathsf{Diff}, \pi_1 \rangle = \langle \mathsf{Diff}, \pi_A \rangle \qquad \text{and} \qquad \langle \mathsf{MonoC}, \pi_1 \rangle = \langle \mathsf{MonoC}, \pi_A \rangle.$$

5.2 General graphs.

Let us start by slightly generalizing concepts of decomposition and decomposable observable.

Definition 21. *A permutation*

$$\mu\colon (\mathbf{x}, \mathbf{y}) \mapsto (\mu_1(\mathbf{x}, \mathbf{y}), \mu_2(\mathbf{x}, \mathbf{y}))$$

*of $S \times S$ is an α-*decomposition *for a local interaction game $\mathcal{G}$ with potential Φ if, for all $(\mathbf{x}, \mathbf{y})$, we have that*

$$|K(\mathbf{x}, \mathbf{y}) - \Phi(\mu_1(\mathbf{x}, \mathbf{y})) - \Phi(\mu_2(\mathbf{x}, \mathbf{y}))| \leqslant \alpha,$$

$\mu_1(\mathbf{x}, \mathbf{y}) = \mu_2(\mathbf{y}, \mathbf{x})$ *and* $\mu_2(\mathbf{x}, \mathbf{y}) = \mu_1(\mathbf{y}, \mathbf{x})$*.*

Note that a decomposition is actually a 0-decomposition (see Definition 14).

Definition 22. *An observable O is α-*decomposable *if it is decomposed by an α-decomposition.*

We next prove that for an α-decomposable observable the extent at which the expectations at stationarity for the one-logit and the all-logit differ depends only on α and β.

Theorem 23. *If observable O is decomposable then*

$$e^{-2\alpha\beta} \cdot \langle O, \pi_1 \rangle \leqslant \langle O, \pi_A \rangle \leqslant e^{2\alpha\beta} \cdot \langle O, \pi_1 \rangle.$$

Proof idea. By mimicking the proof of Theorem 15, we have $e^{-\alpha\beta} Z_1^2 \leqslant Z_A \leqslant e^{\alpha\beta} Z_1^2$ and

$$e^{-\alpha\beta} \sum_{\mathbf{y}} \gamma_1(\mu_1(\mathbf{x}, \mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x}, \mathbf{y})) \leqslant \gamma_A(\mathbf{x}) \leqslant e^{\alpha\beta} \sum_{\mathbf{y}} \gamma_1(\mu_1(\mathbf{x}, \mathbf{y})) \cdot \gamma_1(\mu_2(\mathbf{x}, \mathbf{y}))\,.$$

The theorem then follows by the same arguments of the proof of Theorem 17. □

Finally we prove that for all local interaction games there exists an α-decomposition with α depending only on how far away the social graph G is from being bipartite. Specifically, let us adjust the weights of the social graph such that for each edge e the maximum difference in the potential Φ_e of the two-player game on this edge is exactly 1. Then, we say that a subset of edges of G is *bipartiting* if the removal of these edges makes the graph bipartite. We will denote with $B(G)$ the bipartiting subset of minimum weight and with $b(G)$ its weight. We have then the following theorem.

Theorem 24. *Let $\mathcal{G}$ be a social interaction game on a graph G. Then $\mathcal{G}$ admits an α-decomposition for any $\alpha \geqslant 2 \cdot b(G)$.*

Proof. Let us name as $G' = (V, E')$ the bipartite graph obtained by deleting from G the edges of $B(G)$ and consider the mapping used in the proof of Theorem 19. We know this mapping is actually a permutation and $\mu_1(\mathbf{x}, \mathbf{y}) = \mu_2(\mathbf{y}, \mathbf{x})$ and $\mu_2(\mathbf{x}, \mathbf{y}) = \mu_1(\mathbf{y}, \mathbf{x})$. We will show that, for every $\mathbf{x}, \mathbf{y}$

$$|K(\mathbf{x}, \mathbf{y}) - \Phi(\tilde{\mathbf{x}}) - \Phi(\tilde{\mathbf{y}})| \leqslant 2 \cdot b(G)\,, \tag{15}$$

where $\tilde{\mathbf{x}} = \mu_1(\mathbf{x}, \mathbf{y})$ and $\tilde{\mathbf{y}} = \mu_2(\mathbf{x}, \mathbf{y})$.

Observe that $K(\mathbf{x}, \mathbf{y}) = \sum_{e \in E'} K_e(\mathbf{x}, \mathbf{y}) + \sum_{e \in E \setminus E'} K_e(\mathbf{x}, \mathbf{y})$. From Theorem 19, for each edge $e = (u, v) \in E'$ we have $K_e(\mathbf{x}, \mathbf{y}) = \Phi_e(\tilde{x}_u, \tilde{x}_v) + \Phi_e(\tilde{y}_u, \tilde{y}_v)$. As for each edge $e = (u, v) \in E \setminus E'$ we have that the endpoints are either both at even distance from r or both at odd distance. In both cases, it turns out that

$$\Phi_e(\tilde{x}_u, \tilde{x}_v) + \Phi_e(\tilde{y}_u, \tilde{y}_v) = \Phi_e(x_u, x_v) + \Phi_e(y_u, y_v)\,.$$

Then we distinguish four cases:

1. $u \notin D(\mathbf{x},\mathbf{y})$ and $v \notin D(\mathbf{x},\mathbf{y})$. In this case $K_e(\mathbf{x},\mathbf{y}) = 2 \cdot \Phi_e(x_u, x_v)$ and thus $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(\tilde{x}_u, \tilde{x}_v) + \Phi_e(\tilde{y}_u, \tilde{y}_v)$.
2. $u \in D(\mathbf{x},\mathbf{y})$ and $v \notin D(\mathbf{x},\mathbf{y})$. In this case $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(x_u, x_v) + \Phi_e(y_u, x_v)$ and thus $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(\tilde{x}_u, \tilde{x}_v) + \Phi_e(\tilde{y}_u, \tilde{y}_v)$.
3. $u \notin D(\mathbf{x},\mathbf{y})$ and $v \in D(\mathbf{x},\mathbf{y})$. In this case $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(x_u, x_v) + \Phi_e(x_u, y_v)$ and thus $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(\tilde{x}_u, \tilde{x}_v) + \Phi_e(\tilde{y}_u, \tilde{y}_v)$.
4. $u \in D(\mathbf{x},\mathbf{y})$ and $v \in D(\mathbf{x},\mathbf{y})$. In this case $K_e(\mathbf{x},\mathbf{y}) = \Phi_e(y_u, x_v) + \Phi_e(x_u, y_v)$. Since $|\Phi_e(x_u, x_v) - \Phi_e(y_u, x_v)| \leqslant w_e$ and $|\Phi_e(y_u, y_v) - \Phi_e(x_u, y_v)| \leqslant w_e$, then
$$|K_e(\mathbf{x},\mathbf{y}) - \Phi_e(\tilde{x}_u, \tilde{x}_v) - \Phi_e(\tilde{y}_u, \tilde{y}_v)| \leqslant 2w_e\,.$$

By summing the contribution of every edge we achieve (15). □

6 Mixing time

The all-logit dynamics for a strategic game has the property that, for every pair of profiles $\mathbf{x}, \mathbf{y}$ and for every value of β, the transition probability from $\mathbf{x}$ to $\mathbf{y}$ is strictly positive. In order to give upper bounds on the mixing time, we will use the following simple well-known lemma (see e.g. Theorem 11.5 in [29]).

Lemma 25. *Let P be the transition matrix of an ergodic Markov chain with state space Ω. For every $y \in \Omega$ let us name $\alpha_y = \min\{P(x,y) \,:\, x \in \Omega\}$ and $\alpha = \sum_{y\in\Omega} \alpha_y$. Then the mixing time of P is $t_{\mathrm{mix}} = \mathcal{O}(1/\alpha)$.*

In this section we first give an upper bound holding for every strategic game. We will then focus on a specific local interaction game, the CW-*game*, that is the game theoretic formulation of the Curie-Weiss model in Statistical Physics, and we will give a refined version of the upper bound and a lower bound.

For a strategic game $\mathcal{G}$, in Section 2 we defined the cumulative utility function for the ordered pair of profiles $(\mathbf{x},\mathbf{y})$ as $U(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} u_i(\mathbf{x}_{-i}, y_i)$. Let us name ΔU the size of the range of U,
$$\Delta U = \max\{U(\mathbf{x},\mathbf{y}) \colon \mathbf{x},\mathbf{y} \in S\} - \min\{U(\mathbf{x},\mathbf{y}) \colon \mathbf{x},\mathbf{y} \in S\}\,.$$
By using Lemma 25 we can give a simple upper bound on the mixing time of the all-logit dynamics for $\mathcal{G}$ as a function of β and ΔU.

Theorem 26 (General upper bound)**.** *For any strategic game $\mathcal{G}$ the mixing time of the all-logit dynamics for $\mathcal{G}$ is $\mathcal{O}\left(e^{\beta \Delta U}\right)$.*

Proof. Let P be the transition matrix of the all-logit dynamics for $\mathcal{G}$ and let $\mathbf{x},\mathbf{y} \in S$ be two profiles. From (3) we have that
$$P(\mathbf{x},\mathbf{y}) = \frac{e^{\beta U(\mathbf{x},\mathbf{y})}}{\sum_{\mathbf{z}\in S} e^{\beta U(\mathbf{x},\mathbf{z})}} = \frac{1}{\sum_{\mathbf{z}\in S} e^{\beta (U(\mathbf{x},\mathbf{z}) - U(\mathbf{x},\mathbf{y}))}} \geqslant \frac{1}{|S| e^{\beta \Delta U}}\,.$$
Hence for every $\mathbf{y} \in S$ it holds that
$$\alpha_{\mathbf{y}} \geqslant \frac{e^{-\beta\Delta U}}{|S|}$$
and $\alpha = \sum_{\mathbf{y}\in S} \alpha_{\mathbf{y}} \geqslant e^{-\beta \Delta U}$. The thesis then follows from Lemma 25. □

6.1 Curie-Weiss model

In this section we prove upper and lower bounds on the mixing time of the all-logit dynamics for a special local interaction game, the CW-game. In such a game, every player has two strategies, and we find it convenient to call them $+1$ and -1. The utility of player $i \in [n]$ is the sum of the number of players playing the same strategy as i, minus the number of players playing the opposite strategy; that is, the utility of player $i \in [n]$ at profile $\mathbf{x} = (x_1, \ldots, x_n) \in \{-1, +1\}^n$ is

$$u_i(\mathbf{x}) = x_i \sum_{j \neq i} x_j \,.$$

It is easy to see that such a game is a potential game with potential function

$$\Phi(\mathbf{x}) = - \sum_{\{i,j\} \in \binom{[n]}{2}} x_i x_j \,.$$

Due to the high level of symmetry of the game, the potential of a profile $\mathbf{x}$ depends only on the *number* of players playing ± 1. If we name $k_{\mathbf{x}} := \sum_{i=1}^{n} x_i$ the *magnetization* of $\mathbf{x}$ (notice that the magnetization is the Diff observable discussed in Section 5.1 where strategies were called 0 and 1), we can write the potential of $\mathbf{x}$ as

$$\Phi(\mathbf{x}) = - \frac{k_{\mathbf{x}}^2 - n}{2} \,.$$

The upper bound. Observe that, for the Curie-Weiss model we have $\Delta U = 2n(n-1)$, hence by using Theorem 26 we get directly that

$$t_{\text{mix}} = \mathcal{O}\left(e^{2\beta n(n-1)} \right) \,. \tag{16}$$

Hence it follows that mixing time is $\mathcal{O}(1)$ for $\beta = \mathcal{O}(1/n^2)$ and it is $\mathcal{O}(\text{poly}(n))$ for $\beta = \mathcal{O}(\log n / n^2)$.

In what follows we show that factor "2" at the exponent in (16) can be removed and that a slightly better upper bound can be given for $\beta > \log n / n$.

Lemma 27. *For every* $\mathbf{x}, \mathbf{y} \in \Omega$ *it holds that*

$$P(\mathbf{x}, \mathbf{y}) \geqslant q^{(n+|k_{\mathbf{y}}|)/2} (1-q)^{(n-|k_{\mathbf{y}}|)/2}$$

where

$$q = \frac{1}{1 + e^{2\beta(n-1)}} \,.$$

Proof. Consider a profile $\mathbf{y} \in \{-1, +1\}^n$ and let $k_{\mathbf{y}}$ be its magnetization. Remember that the number of players playing $+1$ and -1 in $\mathbf{y}$ can be written as $\frac{n+k_{\mathbf{y}}}{2}$ and $\frac{n-k_{\mathbf{y}}}{2}$, respectively. If $\mathbf{y}$ has positive magnetization $k_{\mathbf{y}} > 0$, i.e. if the number of players playing $+1$ is larger than the number of players playing -1, then the profile that minimizes $P(\mathbf{x}, \mathbf{y})$ is profile $\mathbf{x}_- = (-1, \ldots, -1)$ where every player plays -1. If we name

$$q = \frac{e^{-\beta(n-1)}}{e^{-\beta(n-1)} + e^{\beta(n-1)}} = \frac{1}{1 + e^{2\beta(n-1)}}$$

the probability that a player in $\mathbf{x}_-$ chooses strategy $+1$ for the next round, we have that

$$P(\mathbf{x}_-, \mathbf{y}) = q^{\frac{n+k_{\mathbf{y}}}{2}} (1-q)^{\frac{n-k_{\mathbf{y}}}{2}} \,.$$

On the other hand, if $\mathbf{y}$ has negative magnetization $k_{\mathbf{y}} < 0$, $P(\mathbf{x}, \mathbf{y})$ is minimized when $\mathbf{x} = \mathbf{x}_+ = (+1, \dots, +1)$ and, since q is also the probability that a player in $\mathbf{x}_+$ chooses strategy -1 for the next round, we have that

$$P(\mathbf{x}_+, \mathbf{y}) = q^{\frac{n-k_{\mathbf{y}}}{2}} (1-q)^{\frac{n+k_{\mathbf{y}}}{2}}$$

and the thesis follows. □

Now we can give an upper bound on the mixing time by using lemmata 25 and 27

Theorem 28 (Upper bound)**.** *The mixing time of the all-logit dynamics for the Curie-Weiss model is*

$$t_{\text{mix}} = \mathcal{O}\left(n e^{\beta n^2}\right) .$$

If $\beta \geqslant \log n / n$ *the mixing time is*

$$t_{\text{mix}} = \mathcal{O}\left(\frac{n e^{\beta n^2}}{2^n}\right) .$$

Proof. From Lemma 27 it follows that for every $\mathbf{y} \in \{-1,+1\}^n$ we have

$$\alpha_{\mathbf{y}} = \min\{P(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in \{-1,+1\}^n\} \geqslant q^{(n+|k_{\mathbf{y}}|)/2} (1-q)^{(n-|k_{\mathbf{y}}|)/2} .$$

Hence

$$\alpha = \sum_{\mathbf{y} \in \{-1,+1\}^n} \alpha_{\mathbf{y}} \geqslant \sum_{\mathbf{y} \in \{-1,+1\}^n} q^{(n+|k_{\mathbf{y}}|)/2} (1-q)^{(n-|k_{\mathbf{y}}|)/2} . \tag{17}$$

Now observe that there are $\binom{n}{\frac{n-k}{2}}$ profiles with magnetization k, and since $q \leqslant 1/2$, the largest terms in (17) are the ones with magnetization as close to zero as possible. In order to give a lower bound to α we will thus consider only terms with magnetization $k = 0$, when n is even, and terms with magnetization $k = \pm 1$, when n is odd.

Case n even: If we consider only profiles with magnetization $k = 0$ in (17) we have that

$$\alpha \geqslant \binom{n}{n/2} [q(1-q)]^{n/2} .$$

By using a standard lower bound for the binomial coefficient (see e.g. Lemma 9.2 in [29]) we have that

$$\binom{n}{n/2} \geqslant \frac{2^n}{n+1} .$$

As for $[q(1-q)]^{n/2}$ we have that

$$\begin{aligned} q(1-q) &= \frac{1}{1+e^{2\beta(n-1)}} \cdot \frac{1}{1+e^{-2\beta(n-1)}} \\ &= \frac{1}{e^{2\beta(n-1)} + 2 + e^{-2\beta(n-1)}} \\ &= \frac{1}{e^{2\beta(n-1)} \left(1 + 2e^{-2\beta(n-1)} + e^{-4\beta(n-1)}\right)} \end{aligned} \tag{18}$$

Now observe that for every $\beta \geqslant 0$ we can bound $1 + 2e^{-2\beta(n-1)} + e^{-4\beta(n-1)} \leqslant 4$. Thus we have that

$$[q(1-q)]^{n/2} \geqslant \frac{1}{2^n e^{\beta n(n-1)}} . \tag{19}$$

Hence

$$\alpha \geqslant \binom{n}{n/2}[q(1-q)]^{n/2} \geqslant \frac{1}{(n+1)e^{\beta n(n-1)}}\,.$$

And by using Lemma 25 we have

$$t_{\text{mix}} = \mathcal{O}\left(ne^{\beta n(n-1)}\right)\,.$$

If β is large enough, say $\beta \geqslant \log n/n$, in (18) we can bound

$$1 + 2e^{-2\beta(n-1)} + e^{-4\beta(n-1)} \leqslant 1 + \frac{1}{n}\,.$$

Thus, in this case we have that

$$[q(1-q)]^{n/2} \geqslant \frac{1}{e^{\beta n(n-1)}\,(1+1/n)^{(n/2)}} \geqslant \frac{1}{e^{\beta n(n-1)}\cdot\sqrt{e}}\,. \tag{20}$$

Hence $\alpha \geqslant \frac{2^n}{(n+1)e^{1/2+\beta n(n-1)}}$ and

$$t_{\text{mix}} = \mathcal{O}\left(\frac{ne^{\beta n(n-1)}}{2^n}\right)\,.$$

Case n odd: If we consider only profiles with magnetization ± 1 in (17) we get

$$\alpha \geqslant 2\binom{n}{\frac{n+1}{2}}q^{\frac{n+1}{2}}(1-q)^{\frac{n-1}{2}} = 2\binom{n}{\frac{n+1}{2}}\sqrt{\frac{q}{1-q}}\,[q(1-q)]^{n/2}\,.$$

Now observe that

$$\sqrt{\frac{q}{1-q}} = e^{-\beta(n-1)} \qquad \text{and} \qquad \binom{n}{\frac{n+1}{2}} \geqslant \frac{1}{2}\cdot\frac{2^n}{n+1}\,.$$

By using bounds (19) and (20) for $[q(1-q)]^{n/2}$ we get $t_{\text{mix}} = \mathcal{O}\left(ne^{\beta(n^2-1)}\right)$ for every $\beta \geqslant 0$ and $t_{\text{mix}} = \mathcal{O}\left(\frac{ne^{\beta(n^2-1)}}{2^n}\right)$ for $\beta \geqslant \log n/n$. □

The lower bound. A key function for the all-logit dynamics for a potential game with potential function Φ is

$$\Upsilon(\mathbf{x},\mathbf{y}) = (n-2)\Phi(\mathbf{x}) + \Psi(\mathbf{x},\mathbf{y})$$

where $\Psi(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n}\Phi(\mathbf{x}_{-i}, y_i)$. Indeed, from (10) it follows that for a "social potential game"

$$\pi(\mathbf{x})P(\mathbf{x},\mathbf{y}) = \frac{1}{Z}e^{\beta((n-2)\Phi(\mathbf{x})+\Psi(\mathbf{x},\mathbf{y}))}\,.$$

In order to give a lower bound on the mixing time, we first show that, for the Curie-Weiss model, $\Upsilon(\mathbf{x},\mathbf{y})$ is symmetric and can be written as a function of the magnetization of the two profiles and the Hamming distance between them.

Lemma 29. *Let $\mathbf{x},\mathbf{y} \in \{-1,+1\}^n$ be two profiles with magnetization $k_{\mathbf{x}}$ and $k_{\mathbf{y}}$ respectively and let $h_{\mathbf{x},\mathbf{y}}$ be their Hamming distance, i.e. the number of players where they differ. Then*

$$(n-2)\Phi(\mathbf{x}) - \Psi(\mathbf{x},\mathbf{y}) = k_{\mathbf{x}}k_{\mathbf{y}} + 2h_{\mathbf{x},\mathbf{y}} - n\,.$$

Proof. We already know that $\Phi(\mathbf{x}) = \frac{n-k_{\mathbf{x}}^2}{2}$. In order to evaluate $\Psi(\mathbf{x},\mathbf{y}) = \sum_{i=1}^n \Phi(\mathbf{x}_{-i}, y_i)$ let us name a, b and c as follows

$$\begin{aligned} a &= \#\{i \in [n] \,:\, x_i = y_i\}\,; \\ b &= \#\{i \in [n] \,:\, x_i = +1,\, y_i = -1\}\,; \\ c &= \#\{i \in [n] \,:\, x_i = -1,\, y_i = +1\}\,. \end{aligned}$$

In other words, a is the number of players playing the same strategy in profiles $\mathbf{x}$ and $\mathbf{y}$, b is the number of players playing $+1$ in $\mathbf{x}$ and -1 in $\mathbf{y}$, and c the number of players playing -1 in $\mathbf{x}$ and $+1$ in $\mathbf{y}$. It holds that

$$\begin{aligned} \Psi(\mathbf{x},\mathbf{y}) &= \sum_{i=1}^n \Phi(\mathbf{x}_{-i}, y_i) \\ &= a\frac{n-k_{\mathbf{x}}^2}{2} + b\frac{n-(k_{\mathbf{x}}-2)^2}{2} + c\frac{n-(k_{\mathbf{x}}+2)^2}{2} \\ &= \frac{1}{2}\left((a+b+c)(n-k_{\mathbf{x}}^2) + 4(b-c)k_{\mathbf{x}} - 4(b+c)\right)\,. \end{aligned} \tag{21}$$

Now observe that $a+b+c = n$, $2(b-c) = k_{\mathbf{x}} - k_{\mathbf{y}}$, and $(b+c) = h_{\mathbf{x},\mathbf{y}}$. Hence from (21) we get

$$\begin{aligned} \Psi(\mathbf{x},\mathbf{y}) &= \frac{1}{2}\left(n(n+k_{\mathbf{x}}^2) + 2(k_{\mathbf{x}} - k_{\mathbf{y}})k_{\mathbf{x}} - 4h_{\mathbf{x},\mathbf{y}}\right) \\ &= \frac{n^2}{2} - \frac{n-2}{2}k_{\mathbf{x}}^2 - k_{\mathbf{x}}k_{\mathbf{y}} - 2h_{\mathbf{x},\mathbf{y}}\,. \end{aligned} \tag{22}$$

Thus

$$(n-2)\Phi(\mathbf{x}) - \Psi(\mathbf{x},\mathbf{y}) = k_{\mathbf{x}}k_{\mathbf{y}} + 2h_{\mathbf{x},\mathbf{y}} - n\,.$$

□

Since the Hamming distance between two profiles is at most n, from the above lemma we get the following observation.

Observation 30. *Let $\mathbf{x}, \mathbf{y}$ be two profiles with $k_{\mathbf{x}}k_{\mathbf{y}} \leqslant 0$, then $(n-2)\Phi(\mathbf{x}) - \Psi(\mathbf{x},\mathbf{y}) \leqslant n$.*

Now we can give a lower bound on the mixing time by using the bottleneck-ratio technique.

Theorem 31 (Lower bound)**.** *The mixing time of the all-logit dynamics for the Curie-Weiss model is*

$$t_{\text{mix}} = \Omega\left(\frac{e^{\beta n(n-2)}}{4^n}\right)\,.$$

Proof. Let $S_- \subseteq \{-1,+1\}^n$ be the set of profiles with negative magnetization

$$S_- = \{\mathbf{x} \in \{-1,+1\}^n \,:\, k_{\mathbf{x}} < 0\}$$

and observe that $\pi(S_-) \leqslant 1/2$. From Observation 30 we have that for every $\mathbf{x} \in S_-$ and $\mathbf{y} \in S_+ = \{-1,+1\}^n \setminus S_-$ it holds that

$$\pi(\mathbf{x})P(\mathbf{x},\mathbf{y}) = \frac{1}{Z}e^{\beta[(n-2)\Phi(\mathbf{x}) - \Psi(\mathbf{x},\mathbf{y})]} \leqslant e^{\beta n}/Z\,. \tag{23}$$

Moreover, if we name $\mathbf{x}_-$ the profile where everyone is playing -1 we have that

$$\pi(S_-) \geqslant \pi(\mathbf{x}_-) \geqslant \frac{1}{Z}e^{-2\beta\Phi(\mathbf{x}_-)} = \frac{1}{Z}e^{\beta n(n-1)}\,. \tag{24}$$

Hence, by using bounds (23) and (24), and the fact that the size of S_- is at most 2^{n-1}, we can bound the bottleneck at S_- with

$$B(S_-) = \frac{Q(S_-, S_+)}{\pi(S_-)} = \frac{\sum_{\mathbf{x} \in S_-} \sum_{\mathbf{y} \in S_+} \pi(\mathbf{x}) P(\mathbf{x}, \mathbf{y})}{\pi(S_-)} \leqslant \frac{2^{2n-2} e^{\beta n}}{e^{\beta n(n-1)}} = \frac{2^{2n-2}}{e^{\beta n(n-2)}} \,.$$

By using the bottleneck-ratio theorem (see e.g. Theorem 7.3 in [26]) it follows that

$$t_{\text{mix}} = \Omega\left(\frac{e^{\beta n(n-2)}}{2^{2n}}\right) \,. \qquad \square$$

Remarks. In this section we proved upper and lower bounds on the mixing time of the all-logit dynamics for the Curie-Weiss model. In particular, the upper bound shows that for $\beta = \mathcal{O}(1/n^2)$ the mixing time is constant and for $\beta = \mathcal{O}(\log n/n^2)$ it is at most polynomial. The lower bound shows that, for every constant $\varepsilon > 0$, if $\beta > (1+\varepsilon)(\log 4)/n$ the mixing time is exponential. When β is between $\Theta(\log n/n^2)$ and $\Theta(1/n)$ we still cannot say if mixing is polynomial or exponential.

7 Conclusions and open problems

In this paper we considered the selection rule that assigns positive probability only to the set of all players. A natural extension of this selection rule assigns a different probability to each subset of the players. What is the impact of such a probabilistic selection rule on reversibility and on observables? Some interesting results along that direction have been obtained in [1, 2]. Notice that if we consider the selection rule that selects player i with probability $p_i > 0$ (the one-logit sets $p_i = 1/n$ for all i) then the stationary distribution is the same as the stationary distribution of the one-logit. Therefore, all observables have the same expected value and all potential games are reversible.

It is a classical result that the Gibbs distribution that is the stationary distribution of the one-logit (the micro-canonical ensemble, in Statistical Mechanics parlance) is the distribution that maximizes the entropy among all the distributions with a fixed average potential. Can we say something similar for the stationary distribution of the all-logit? A promising direction along this line of research is suggested by results in Section 5: at least in some cases the stationary distribution of the all-logit can be seen as a composition of more simple distributions.

References

[1] Carlos Alós-Ferrer and Nick Netzer. The logit-response dynamics. *Games and Economic Behavior*, 68:413–427, 2010.

[2] Carlos Alós-Ferrer and Nick Netzer. Robust stochastic stability. Working Paper No. 63, Department of Economics, University of Zurich, 2012.

[3] Elliot Anshelevich, Anirban Dasgupta, Éva Tardos, and Tom Wexler. Near-optimal network design with selfish agents. In *STOC*, pages 511–520, 2003.

[4] Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, Paolo Penna, and Giuseppe Persiano. Convergence to equilibrium of logit dynamics for strategic games. In *Proc. of the 23rd ACM Symp. on Parallelism in Algorithms and Architectures (SPAA'11)*, pages 197–206. ACM, 2011.

[5] Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, and Giuseppe Persiano. Metastability of logit dynamics for coordination games. In *Proc. of the ACM-SIAM Symp. on Discrete Algorithms (SODA'12)*, pages 1006–1024. SIAM, 2012.

[6] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1230, September 2000.

[7] Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131:311–340, 2005. Preliminary version in FOCS 01.

[8] David Bindel, Jon M. Kleinberg, and Sigal Oren. How bad is forming your own opinion? In *FOCS*, 2011.

[9] Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5:387–424, 1993.

[10] Christian Borgs, Jennifer T. Chayes, Jian Ding, and Brendan Lucier. The hitchhiker's guide to affiliation networks: A game-theoretic approach. In *ICS*, pages 389–400, 2011.

[11] Christian Borgs, Jennifer T. Chayes, Brian Karrer, Brendan Meeder, R. Ravi, Ray Reagans, and Amin Sayedi. Game-theoretic models of information overload in social networks. In *WAW*, pages 146–161, 2010.

[12] Jacomo Corbo and David C. Parkes. The price of selfish behavior in bilateral network formation. In *PODC*, pages 99–107, 2005.

[13] A. Cournot. *Recherches sur le Principes mathematiques de la Theorie des Richesses*. Hachett, Paris, 1838.

[14] Glenn Ellison. Learning, local interaction, and coordination. *Econometrica*, 61(5):1047–1071, 1993.

[15] Alex Fabrikant, Ankur Luthra, Elitza N. Maneva, Christos H. Papadimitriou, and Scott Shenker. On a network creation game. In *PODC*, pages 347–351, 2003.

[16] Diodato Ferraioli, Paul Goldberg, and Carmine Ventre. Decentralized dynamics for finite opinion games. In *Algorithmic Game Theory (SAGT '12)*, pages 144–155. Springer Berlin / Heidelberg, 2012.

[17] Drew Fudenberg and David K. Levine. *The Theory of Learning in Games.* MIT, 1998.

[18] Drew Fudenberg and Jean Tirole. *Game Theory.* MIT, 1992.

[19] John C. Harsanyi and Reinhard Selten. *A General Theory of Equilibrium Selection in Games.* MIT Press, 1988.

[20] Sergiu Hart and Andreu Mas-Colell. A general class of adaptive procedures. *J. Econ. Theory*, 98:26–54, 2001.

[21] Matthew O. Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74, October 1996.

[22] Frank Kelly. *Reversibility and Stochastic Networks.* Wiley, 1979.

[23] Jon M. Kleinberg and Sigal Oren. Mechanisms for (mis)allocating scientific credit. In *STOC*, pages 529–538, 2011.

[24] L. D. Landau and E. M. Lifshitz. *Statistical Physics, 3rd Edition Part 1.* Butterworth-Heinemann, Oxford, 1996.

[25] David Levin, Malwina Luczak, and Yuval Peres. Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probability Theory and Related Fields*, 146:223–265, 2010.

[26] David Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times.* American Mathematical Society, 2008.

[27] Fabio Martinelli. Lectures on Glauber dynamics for discrete spin models. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Math.*, pages 93–191. Springer, 1999.

[28] Daniel L. McFadden. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, pages 105–142. Academic Press: New York, 1974.

[29] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis.* Cambridge University Press, 2005.

[30] Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[31] Andrea Montanari and Amin Saberi. Convergence to equilibrium in local interaction games. In *Proc. of the 50th Ann. Symp. on Foundations of Computer Science (FOCS'09)*. IEEE, 2009.

[32] Luca Trevisan. Max cut and the smallest eigenvalue. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 263–272. ACM, 2009.

[33] David H. Wolpert. Information theory – the bridge connecting bounded rational game theory and statistical physics. In *Complex Engineered Systems*, volume 14, pages 262–290. Springer Berlin / Heidelberg, 2006.

[34] H. Peyton Young. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions.* Princeton University Press, 1998.

[35] H. Peyton Young. The diffusion of innovations in social networks. Economics Working Paper Archive number 437, Johns Hopkins University, Department of Economics, 2000.

A A characterization of potential games

In this section we review a characterization of potential games in terms of the utilities. Let $\mathcal{G}$ be a game. A *circuit* $\Gamma = \langle s_0, \ldots, s_\ell \rangle$ is a sequence of strategy profiles such that $s_0 = s_\ell$, $s_h \neq s_k$ for $1 \leqslant h \neq k \leqslant \ell$ and, for $k = 1, \ldots, \ell$, there exists player i_k such that s_{k-1} and s_k differ only for player i_k. For such a circuit Γ we define the *utility improvement* $I(\Gamma)$ as

$$I(\Gamma) = \sum_{k=1}^{\ell} \left[u_{i_k}(s_k) - u_{i_k}(s_{k-1}) \right] .$$

The following theorem holds.

Theorem 32 (Monderer and Shapley [30]). *A game $\mathcal{G}$ is a potential game if and only if $I(\Gamma) = 0$ for all circuits of length 4.*