A CLASS OF MULTI-PARTICLE REINFORCED INTERACTING RANDOM WALKS

JUN CHEN

ABSTRACT. We consider a class of multi-particle reinforced interacting random walks. In this model, there are some (finite or infinite) particles performing random walks on a given (finite or infinite) connected graph, so that each particle has higher probability to visit neighboring vertices or edges which have been seldom visited by the other particles. Specifically we investigate two particles' vertex-reinforced interacting random walks on finite complete graphs. By a dynamical approach we prove that the two particles' occupation measure asymptotically has small joint support almost surely if the reinforcement is strong.

1. INTRODUCTION

"Reinforced random walk" (RRW) is a class of non-Markovian random walk, which has been extensively studied in the last twenty years, see the survey [16]. RRW is defined on the vertices of an undirected graph in such a way that the probability of a transition from one vertex to another depends on the number of previous transitions along the neighboring edges (respectively, vertices). We refer to such models as “edge-reinforced random walk” (ERRW) (respectively, “vertex-reinforced random walk” (VRRW)).

RRWs can be defined on all kinds of graphs, either finite or infinite. It also can be defined by very different reinforcement mechanisms through the weight functions, which could lead to quite different phenomena to occur. A question of interest is the recurrence or transience of RRW. For example, Angel et al. [1] and Sabot et al. [19] independently showed that linearly ERRW on any graph with bounded degrees is recurrent. Another phenomenon of interest is localization. For example, Pemantle and Volkov [17] showed that VRRW on \( \mathbb{Z} \) has finite range, and then Tarrés [21] showed that VRRW on \( \mathbb{Z} \) eventually gets stuck on five points. Phase transitions are also of interest. For instance, Pemantle [14] showed that ERRW on infinite binary tree can vary from transient to recurrent, depending on the value of an adjustable parameter measuring the strength of the feedback (reinforcement), and Volkov [22] showed that VRRW on \( \mathbb{Z} \) with weight function \( k^\alpha \) will just visit 2, 5, \( \infty \) sites after a large time \( T_0 \) when \( \alpha > 1, = 1 \) or \( < 1 \) respectively.
So far, almost all the considered RRW models are one particle’s self-interacting with positive (attractive) feedback (reinforcement), i.e. the edges (or vertices) already being visited more are more likely to be visited in the future. In analyzing such models, four main methods are commonly used: exchangeability \([9,16]\), branching process embedding \([5,11,12,20]\), stochastic approximation via martingale methods \([7,18]\), and dynamical system approach \([2,3]\).

One direction to generalize RRW is to consider multi-particle RRW. Kovchegov \([9]\) showed for the two particles’ edge reinforced process on \(\mathbb{Z}\), the two particles meet infinitely often a.s.. In his model, each particle’s transition probability is positively reinforced (determined) by both particles’ visits on the edges. Except for his paper, the generalization of RRW to multi-particle RRW models with more general reinforced interacting mechanism hasn’t appeared yet. In 2010, Itai Benjamini proposed a class of new multi-particle reinforced interacting random walks, and finally our paper is an exploration of that model.

2. The model and statement of result

We will define a class of two particles’ vertex-reinforced interacting random walks on a connected graph. Denote the two particles by \(X\) and \(Y\), and the graph by \(G = (V, E)\). At each step, both \(X\) and \(Y\) will jump to some vertices on \(V\). Let \(X_k, Y_k\) be \(X, Y\)’s locations at time \(k\) on \(V\), and \(N(X, v, n), N(Y, v, n)\) be the number of \(X, Y\)’s visits to vertex \(v\) by time \(n\). We also assume that \(N(X, v, 0) = N(Y, v, 0) = 1\) for any \(v \in V\). Denote the natural filtration generated by \(\{X_k, 0 \leq k \leq n\}\) and \(\{Y_k, 0 \leq k \leq n\}\) by \(\mathcal{F}_n (n \in \mathbb{N})\). Then the stochastic process \((X_n, Y_n)\)’s transition probability is defined

\[
P(X_{n+1} = v | \mathcal{F}_n) = \frac{1_{v \sim X_n} \cdot w_{N(Y,v,n)}}{\sum_{p \sim X_n} w_{N(Y,p,n)}}
\]

and

\[
P(Y_{n+1} = v | \mathcal{F}_n) = \frac{1_{v \sim Y_n} \cdot w_{N(X,v,n)}}{\sum_{p \sim Y_n} w_{N(X,p,n)}},
\]

where "\(\sim\)" represents some vertex relation on \(G\) (e.g. nearest-neighbor), and \(w_k (k \in \mathbb{Z}_+)\) is a fixed sequence of positive numbers, referred to as “weights”. One natural weight sequence to work with is the nonlinear one \(w_k = k^{-\alpha}\) for some \(\alpha > 0\).

On a finite connected graph \(G = (V, E)\) with \(V = \{1, \ldots, d\}\), one can set

\[
x_i(n) = \frac{N(X, i, n)}{n + d}, \quad y_i(n) = \frac{N(Y, i, n)}{n + d}, \quad \forall i \in V
\]

as \(X\) and \(Y\)’s empirical occupation measure on \(V\) by time \(n\). Then if \(w_k\) (like \(k^{-\alpha}\)) is homogenous in \(k\), we can rewrite \((2.1)\) and \((2.2)\) as

\[
P(X_{n+1} = i | \mathcal{F}_n) = \frac{1_{i \sim X_n} \cdot \varphi(y_i(n))}{\sum_{j \sim X_n} \varphi(y_j(n))}
\]

and

\[
P(Y_{n+1} = i | \mathcal{F}_n) = \frac{1_{i \sim Y_n} \cdot \varphi(x_i(n))}{\sum_{j \sim Y_n} \varphi(x_j(n))},
\]

where \(\varphi(x)\) is a function \(\varphi : [0,1] \to \mathbb{R}_{\geq 0}\). Hence, \(X, Y\)’s transition probabilities are defined by their occupation measure.
In this paper, we will work with the model defined by (2.4) and (2.5) on finite complete graphs with \( \varphi(x) = [\delta 1_{x \leq \delta} + x 1_{x > \delta}]^{-\alpha} \) for some fixed \( \delta, \alpha > 0 \). Here we assume \( \delta \) is small enough and the vertex relation “\(~\)” is that the graph distance between two vertices is less than or equal to 1, i.e. \( X, Y \) are lazy random walks. Then for any \( i, j \in V \), (2.4) and (2.5) become

\[
\mathbb{P}(X_{n+1} = i | \mathcal{F}_n) = \frac{[\delta 1_{x_i(n) \leq \delta} + y_i(n) 1_{y_i(n) > \delta}]^{-\alpha}}{\sum_{k=1}^d [\delta 1_{y_k(n) \leq \delta} + y_k(n) 1_{y_k(n) > \delta}]^{-\alpha}}
\]

(2.6)

and

\[
\mathbb{P}(Y_{n+1} = j | \mathcal{F}_n) = \frac{[\delta 1_{x_j(n) \leq \delta} + y_j(n) 1_{y_j(n) > \delta}]^{-\alpha}}{\sum_{k=1}^d [\delta 1_{x_k(n) \leq \delta} + x_k(n) 1_{x_k(n) > \delta}]^{-\alpha}}.
\]

(2.7)

Denote \( x(n) = (x_1(n), \ldots, x_d(n)) \) and \( y(n) = (y_1(n), \ldots, y_d(n)) \). Set a 2d dimensional vector

\[
z(n) = (x(n), y(n)) = (x_1(n), \ldots, x_d(n), y_1(n), \ldots, y_d(n)).
\]

(2.8)

Notice that \( z(n) \) is a Markov chain living in \( \mathbb{R}^{2d} \). Then we will show \( z(n) \)'s asymptotic behavior.

Here we want to mention that when \( \min_{i,j \in V} \{x_i(n), y_j(n)\} > \delta \), (2.6) and (2.7) are equivalent to the following formulas

\[
\mathbb{P}(X_{n+1} = i | \mathcal{F}_n) = \frac{N(Y, i, n)^{-\alpha}}{\sum_{k=1}^d N(Y, k, n)^{-\alpha}}, \quad \forall i \in V
\]

(2.9)

\[
\mathbb{P}(Y_{n+1} = j | \mathcal{F}_n) = \frac{N(X, j, n)^{-\alpha}}{\sum_{k=1}^d N(X, k, n)^{-\alpha}}, \quad \forall j \in V,
\]

(2.10)

which can be easily seen as a class of two particles’ repelling interacting random walks with nonlinear reinforcement.

Throughout the paper, we use the following notation.

**Notation 2.1.**

1. Let \( \Delta \) be the closed \((d - 1)\)-dimensional simplex

\[
\Delta = \left\{ u \in \mathbb{R}^d : u_i \geq 0, \sum_{i \in V} u_i = 1 \right\}.
\]

Denote the relative interior of \( \Delta \) by \( \overset{\circ}{\Delta} \).

2. Let \( D \) be the product of two simplices \( \Delta \times \Delta \)

\[
D = \left\{ (u, v) \in \mathbb{R}^{2d} : u_i \geq 0, \sum_{i \in V} u_i = 1 \text{ and } v_i \geq 0, \sum_{i \in V} v_i = 1 \right\}.
\]

Denote the relative interior of \( D \) by \( \overset{\circ}{D} \), and the boundary of \( D \) by \( \partial D \).

3. Let \( TD \) be the set identified with the tangent space to \( D \) at each point

\[
TD = T(\Delta \times \Delta) = \left\{ (u, v) \in \mathbb{R}^{2d}, \sum_{i \in V} u_i = 0, \sum_{i \in V} v_i = 0 \right\}.
\]

4. Let \( U \) be the \( d \)-dimensional vector \((1/d, \ldots, 1/d)\). We also call \( U \) the uniform distribution.

5. Let \( \| \cdot \| \) denote the \( L^1 \) norm on \( \mathbb{R}^{2d} \) defined by \( \| v \| = \sum_{i=1}^{2d} |v_i| \).
Then we have the following theorem for any small positive $\delta$ appearing in the definition of $\varphi(x)$.

**Theorem 2.2.** For any fixed $d \geq 3 \in \mathbb{N}$, there exists some $\alpha(d)$ independent of $\delta$, s.t. when $\alpha \geq \alpha(d)$, the two components $x(n)$ and $y(n)$ of $z(n)$ in (2.8) asymptotically have small joint support bounded by $4\delta$ almost surely, i.e.

$$
\mathbb{P} \left\{ \exists n_0, \bigcap_{n \geq n_0} \left\{ \sum_{i=1}^{d} x_i(n)y_i(n) < 4\delta \right\} \right\} = 1.
$$

The organization of the rest of this paper is as follows: In Section 3, we will show that $z(n)$ belongs to a class of stochastic approximation algorithms. In Section 4, we will introduce the dynamical approach and conclude that the limit set of $z(n)$ is contained in the chain recurrent set of a semiflow induced by an ordinary differential equation (ODE). In Section 5, we will prove Theorem 2.2. In Section 6, some further open problems are proposed.

### 3. Stochastic approximation algorithms

In general, a **stochastic approximation algorithm** is a discrete time stochastic process whose form can be written as

$$
z(n+1) - z(n) = \gamma_n H(z(n), \xi(n))$$

where $H : \mathbb{R}^m \times \mathbb{R}^k \rightarrow \mathbb{R}^m$ is a measurable function that characterizes the algorithm, $\{z(n)\}_{n \geq 0} \subset \mathbb{R}^m$ is the sequence of parameters to be recursively updated, $\{\xi(n)\}_{n \geq 0} \subset \mathbb{R}^k$ is a sequence of random variables defined on some probability space $(\Omega, \mathcal{G}, \mathbb{P})$, and $\{\gamma_n\}_{n \geq 0}$ is a sequence of “small” nonnegative numbers. Such processes were first introduced in the early 50s in the works of Robbins and Monro [18] and Kiefer and Wolfowitz [8].

To show that $z(n)$ in (2.8) is a stochastic approximation algorithm, we need to show $z(n)$ satisfies a difference equation of the form (3.1). Observe that from (2.3)

$$
x_i(n+1) - x_i(n) = \frac{N(X,i,n) + 1_{X_{n+1}=i}}{n+1+d} - \frac{N(X,i,n)}{n+d},
$$

(3.2)

Similarly, we can derive a difference equation for $y_i(n)$. So $z(n)$ satisfies (3.1) with

$$
\gamma_n = \frac{1}{n+1+d}, \quad \xi(n) = (1_{X_{n+1}=1}, \ldots, 1_{X_{n+1}=d}, 1_{Y_{n+1}=1}, \ldots, 1_{Y_{n+1}=d})
$$

and $H : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$ defined by

$$
H(z(n), \xi(n)) = -z(n) + \xi(n).
$$

That is,

$$
z(n+1) - z(n) = \frac{1}{n+1+d} (-z(n) + \xi(n)).
$$

(3.4)

To analyze the asymptotic behavior of $z(n)$ in (3.4), it is convenient to introduce a related ODE as shown in the next section.
4. The dynamical approach

The dynamical approach is a method used to analyze stochastic approximations, introduced by Ljung [13] and Kushner and Clark [10]. The idea is to decouple the stochastic approximation algorithm into its mean part and the other so-called “noise” part, and then study the asymptotic behavior of the algorithm in terms of the mean component’s behavior. This method has been widely studied and inspired many works, such as the book by Kushner and Clark [10], numerous articles by Kushner, and more recently the book by Benveniste, Metivier, and Priouret [4].

In the above perspective, our stochastic approximation algorithm can be written as

\[ z(n+1) - z(n) = \gamma_n \{ ((-z(n)) + \mathbb{E}[\xi(n) | \mathcal{F}_n]) + (\xi(n) - \mathbb{E}[\xi(n) | \mathcal{F}_n]) \}, \]

where \( \mathcal{F}_n \) was defined in Section 2. Set a map \( \pi = (\pi_1, \ldots, \pi_d) : \Delta \rightarrow \Delta \) with

\[ \pi_i(x) = \frac{\varpi(x_i)}{\sum_{k=1}^{d} \varpi(x_k)}, \quad \forall x \in \Delta, \]

where \( \varpi(x) = [\delta_{1 \leq x} + x(1-x)]^{-\alpha} \). Observe that, by (2.6), (2.7) and (3.3),

\[ \mathbb{E}[\xi(n) | \mathcal{F}_n] = (\pi(y(n)), \pi(x(n))). \]

Thus, defining \( \{ u_n \}_{n \geq 0} \subset \mathbb{R}^{2d} \) by

\[ u_n = \xi(n) - \mathbb{E}[\xi(n) | \mathcal{F}_n] \]

and \( F = (F_1, \ldots, F_{2d}) \) to be a vector field in \( \Delta \times \Delta \) with

\[ F_i(x_1, \ldots, x_d, y_1, \ldots, y_d) = \begin{cases} -x_i + \pi_i(y_1, \ldots, y_d), & \text{if } 1 \leq i \leq d; \\ -y_{i-d} + \pi_{i-d}(x_1, \ldots, x_d), & \text{if } d+1 \leq i \leq 2d, \end{cases} \]

our random process takes the form

\[ z(n+1) - z(n) = \gamma_n [F(z(n)) + u_n]. \]

The above expression is a particular case of a class of stochastic approximation algorithms studied by Benaim in [2], on which he related the behavior of the algorithm to a weak notion of recurrence for the ODE: that of chain-recurrence. His theorem asserts that, under the assumptions of Kushner and Clark lemma [10], the accumulation points of \( \{ z(n) \}_{n \geq 0} \) are contained in the chain-recurrent set of the semiflow generated by the ODE.

In the remaining of this section, we introduce the necessary definitions for semiflows, then state Benaim’s theorem, and conclude the section by proving that our model satisfies the required conditions of this theorem.

4.1. Preliminaries on semiflows. Let \( \Gamma \subset \mathbb{R}^m \) be a metric space and “dist(·,·)” denote the metric. Let \( \Phi : \mathbb{R}_{\geq 0} \times \Gamma \rightarrow \Gamma \) be a continuous map. For simplicity, denote \( \Phi(t, x) \) by \( \Phi_t(x) \).

**Definition 4.1** (Semiflow). A semiflow on \( \Gamma \) is a continuous map \( \Phi : \mathbb{R}_{\geq 0} \times \Gamma \rightarrow \Gamma \) such that

(i) \( \Phi_0 \) is the identity on \( \Gamma \), and

(ii) \( \Phi_{t+s} = \Phi_t \circ \Phi_s \) for any \( t, s \geq 0 \).
In particular, for every continuous vector field $F: \mathbb{R}^m \to \mathbb{R}^m$ with unique integral curves, we can associate a semiflow on $\mathbb{R}^m$ by the equation
\[
\frac{d}{dt}\Phi_t(x) = F(\Phi_t(x)), \quad \forall x \in \mathbb{R}^m, \forall t \in \mathbb{R}_{\geq 0}.
\]
If $F$ is Lipschitz, it has unique integral curves.

Fix a semiflow $\Phi$ on $\Gamma \subset \mathbb{R}^m$.

**Definition 4.2** (Invariant set). A set $A \subset \Gamma$ is called invariant if $\Phi_t(A) \subset A$ for every $t \geq 0$.

**Definition 4.3** (Equilibrium point). A point $x \in \Gamma$ is called an equilibrium if $\Phi_t(x) = x$ for all $t \geq 0$. The equilibrium set of $\Phi$ is the set of all equilibrium points.

When $\Phi$ is induced by a vector field $F$, the equilibrium set coincides with the set on which $F$ vanishes.

**Definition 4.4** (Chain-recurrent point). Given $\rho, T > 0$, a point $x \in \Gamma$ is called $(\rho, T)$-recurrent if there are points $x_0 = x, x_1, \ldots, x_{k-1}, x_k = x \in \Gamma$ and real numbers $t_0, t_1, \ldots, t_{k-1} \geq T$ such that
\[
\text{dist}(\Phi_{t_i}(x_i), x_{i+1}) < \rho, \quad i = 0, \ldots, k-1.
\]
x is said to be chain-recurrent if it is $(\rho, T)$-recurrent for any $\rho, T > 0$.

We denote by $CR(\Phi)$ the set of chain-recurrent points. $CR(\Phi)$ is closed and invariant.

We denote the limit set of a discrete sequence $\{x(n)\}_{n \geq 0} \subset \Gamma$ by $L(\{x(n)\}_{n \geq 0})$. The sets describing the asymptotic behavior of the orbits of $\Phi$ are the omega limit sets.

**Definition 4.5** (Omega limit set). The omega limit set of $w \in \Gamma$, denoted by $\omega(w)$, is the set of $x \in \Gamma$ such that $\lim_{k \to \infty} \Phi_{t_k}(w) = x$ for some sequence $t_k > 0$ with $\lim_{k \to \infty} t_k = \infty$.

If $\Gamma$ is compact, $\omega(w)$ is a nonempty, compact, connected and invariant set.

**Definition 4.6** (Lyapunov function). A continuous map $L: \Gamma \to \mathbb{R}$ is said to be a Lyapunov function for some subset $\Lambda \subset \Gamma$ if the function $t \in \mathbb{R}_{\geq 0} \to L(\Phi_t(x))$ is strictly decreasing along any non-constant orbit $\Phi_t(x) \subset \Lambda$.

4.2. A limit set theorem. The reason we can characterize the limit set of the random process via the chain-recurrent set of the deterministic semiflow is due to Theorem 1.2 of [2] which, to our purposes, is stated as

**Theorem 4.7.** Let $F: \mathbb{R}^m \to \mathbb{R}^m$ be a continuous vector field with unique integral curves, and let $\{z(n)\}_{n \geq 0}$ be a solution to the recursion
\[
z(n+1) - z(n) = \gamma_n [F(z(n)) + u_n],
\]
where $\{\gamma_n\}_{n \geq 0}$ is a decreasing gain sequence\(^1\) and $\{u_n\}_{n \geq 0} \subset \mathbb{R}^m$. Assume that
(i) $\{z(n)\}_{n \geq 0}$ is bounded, and
(ii) for each $T > 0$,
\[
\lim_{n \to \infty} \sup_k \left\{ \left\| \sum_{i=n}^{k-1} \gamma_i u_i \right\| : \sum_{i=n}^{k-1} \gamma_i \leq T \right\} = 0.
\]

\(^1\) $\lim_{n \to \infty} \gamma_n = 0$ and $\sum_{n \geq 0} \gamma_n = \infty$. 
Hence, \( L(\{z(n)\}_{n \geq 0}) \) is a connected set chain-recurrent for the semiflow induced by \( F \).

4.3. The random process \((4.4)\) satisfies Theorem \(4.7\) First, note that \( \delta I_{x \leq \delta} + x I_{x > \delta} \) is bounded by \( \delta \) and 1, and \( \varpi(x) \) is Lipschitz. Then \( \pi \) in \((4.1)\) and \( F \) in \((4.3)\) are Lipschitz. Meanwhile, \( \gamma_n = 1/(n + 1 + d) \) satisfies

\[
\lim_{n \to \infty} \gamma_n = 0 \quad \text{and} \quad \sum_{n \geq 0} \gamma_n = \infty.
\]

It remains to check condition (ii). For that, let \( M_n = \sum_{i=0}^{n} \gamma_i u_i \). Observe that \( \{M_n\}_{n \geq 0} \) is a martingale adapted to \( \{F_{n+1}\}_{n \geq 0} \)

\[
\mathbb{E}[M_{n+1}|F_{n+1}] = \sum_{i=0}^{n} \gamma_i u_i + \mathbb{E}[\gamma_{n+1} u_{n+1}|F_{n+1}] = \sum_{i=0}^{n} \gamma_i u_i = M_n.
\]

Furthermore, because for any \( n \geq 0 \)

\[
\sum_{i=0}^{n} \mathbb{E}[\|M_{i+1} - M_i\|^2|F_{i+1}] \leq (2d)^2 \sum_{i=0}^{n} \gamma_i^2 \leq (2d)^2 \sum_{i=0}^{n} \gamma_i^2 < \infty \quad \text{a.s.,}
\]

the sequence \( \{M_n\}_{n \geq 0} \) converges to a finite random variable in \( \mathbb{R}^{2d} \) almost surely (see e.g. Theorem 5.4.9 of [6]). In particular, it is a Cauchy sequence and so condition (ii) holds almost surely.

Now, in view of Theorem \(4.7\) we will investigate the chain-recurrent set of semiflow generated by the ODE

\[
\begin{aligned}
\frac{du_i(t)}{dt} &= -u_i(t) + \frac{f(u_i)^{-\alpha}}{\sum_{k=1}^{d} f(v_k)^{-\alpha}}, \quad i = 1, \ldots, d \\
\frac{dv_i(t)}{dt} &= -v_i(t) + \frac{f(v_i)^{-\alpha}}{\sum_{k=1}^{d} f(u_k)^{-\alpha}}, \quad i = 1, \ldots, d
\end{aligned}
\]  

(4.5)

where

\[
f(x) = \delta I_{x \leq \delta} + x I_{x > \delta}, \quad \text{i.e.} \ f(x) = \varpi(x)^{-\frac{1}{\alpha}}.
\]

(4.6)

We can rewrite (4.5) in vector form

\[
\begin{aligned}
\frac{du(t)}{dt} &= -u(t) + \pi(v(t)) \\
\frac{dv(t)}{dt} &= -v(t) + \pi(u(t))
\end{aligned}
\]  

or \( \frac{d\Xi(t)}{dt} = F(\Xi(t)) \)

(4.7)

where \( \Xi(t) = (u(t), v(t)) \in D \).

Before moving to the proof of Theorem \(2.2\) we will prove a simple fact regarding \(4.6\).

**Proposition 4.8.** The domain \( D \) is invariant under \( \Phi \), the semiflow induced by \(4.5\).

**Proof.** Suppose \( (u, v) \in \partial D \). Without loss of generality, we can assume that there exists some \( i \in \mathbb{V} \) such that \( u_i = 0 \). Then by \((4.3)\), we have

\[
\left. \frac{du_i(t)}{dt} \right|_{(u,v)} \geq \inf_{v \in \Delta} \frac{f(v_i)^{-\alpha}}{\sum_{j=1}^{d} f(v_j)^{-\alpha}} > 0.
\]

Hence, \( F(u, v) \) points inward whenever \( (u, v) \) belongs to the boundary of \( D \). Thus any forward trajectory based in \( D \) remains in \( D \). \( \Box \)
5. Proof of Theorem 2.2

According to Theorem 4.7, the limit set of \( \{z(n)\}_{n \geq 0} \) is contained in the chain recurrent set, and so the first step to prove Theorem 2.2 is to characterize chain recurrent set for our specific semiflow induced by (4.5). Recall \( U \)'s definition in Notation 2.1. We will conclude the proof of Theorem 2.2 by showing that \( \{z(n)\}_{n \geq 0} \) has probability 0 to converge to the isolated unstable equilibrium \((U, U)\).

5.1. Chain recurrent set.

5.1.1. Lyapunov function. We characterize our chain-recurrent set \( \text{CR}(\Phi) \) by introducing a Lyapunov function

\[
L(u, v) = \sum_{i=1}^{d} u_i v_i, \quad (u, v) \in D. \quad (5.1)
\]

Let \( \Phi_t(x) = (u_1(t), \ldots, u_d(t), v_1(t), \ldots, v_d(t)) \) \((t \geq 0)\) be an orbit of \( \Phi \) where \( x = (u_1(0), \ldots, u_d(0), v_1(0), \ldots, v_d(0)) \). Then

\[
\frac{d}{dt}(L(\Phi_t(x))) = \sum_{i=1}^{d} v_i(t) \frac{du_i(t)}{dt} + \sum_{i=1}^{d} u_i(t) \frac{dv_i(t)}{dt} = \sum_{i=1}^{d} v_i \left(-u_i + \frac{f(v_i) - \alpha}{\sum_{k=1}^{d} f(v_k) - \alpha}\right) + \sum_{i=1}^{d} u_i \left(-v_i + \frac{f(u_i) - \alpha}{\sum_{k=1}^{d} f(u_k) - \alpha}\right) = -2 \sum_{i=1}^{d} u_i v_i + \sum_{i=1}^{d} \frac{u_i f(u_i) - \alpha}{f(v_i) - \alpha} + \sum_{i=1}^{d} \frac{v_i f(v_i) - \alpha}{f(u_i) - \alpha}. \quad (5.2)
\]

Notice that the right hand side of (5.2) depends on \( t \) only through dependence on \( u_i(t) \) and \( v_i(t) \). We have the following lemma about (5.2), which verifies that \( L(u, v) \) is a Lyapunov function for a large subset of the domain \( D \) according to Definition 4.6.

**Lemma 5.1.** Let \( D^\delta = \{(u, v) \in D : L(u, v) \geq 3\delta\} \). For any fixed \( d \geq 3 \in \mathbb{N} \), there exists some \( \alpha(d) \) independent of \( \delta \), s.t. when \( \alpha \geq \alpha(d) \)

\[
\left. \frac{d}{dt}(L(\Phi_t(x))) \right|_{(u,v)} \leq 0, \quad \forall (u, v) \in D^\delta, \quad (5.3)
\]

with equality if and only if \( (u, v) = (U, U) \).

To prove this lemma, we need several other lemmas. Recall that \( V = \{1, \ldots, d\} \).

**Lemma 5.2.** When \( \alpha > d - 2 \), \( U \) (uniform distribution) is a local minimum in \( \Delta \) of the following function

\[
g(u_1, \ldots, u_d) := 2 \min_{i \in V} u_i - d \left( \min_{i \in V} u_i \right)^2 - \frac{\sum_{i=1}^{d} u_i - \alpha}{\sum_{i=1}^{d} u_i - (\alpha + 1)}. \]

In particular, \( g\left(\frac{1}{d}, \ldots, \frac{1}{d}\right) = 0 \).
Proof. Define a function on \((\mathbb{R}^d)^+\):

\[
G(w_1, \ldots, w_d) = \frac{2}{\sum_{i=1}^d w_i} \cdot \min_{i \in V} w_i - \frac{d}{(\sum_{i=1}^d w_i)^2} \cdot (\min_{i \in V} w_i)^2 - \frac{1}{\sum_{i=1}^d w_i} \cdot \frac{\sum_{i=1}^d w_i^{-\alpha}}{\sum_{i=1}^d w_i^{-(\alpha+1)}}.
\]

Observe that \(G(w_1, \ldots, w_d)\) is a homogeneous function, and that it has the same value as \(g(u_1, \ldots, u_d)\) whenever

\[
u_i = \frac{w_i}{\sum_{j=1}^d w_j}, \forall i \in V.
\]

Let \(W = (w, \ldots, w) (w > 0)\) (we refer to \(W\) as the diagonal). So to prove the lemma, it suffices to prove that \(W\) is a local minimum of \(G(w_1, \ldots, w_d)\).

Without loss of generality, we can assume \(w_d = \min_{i \in V} w_i\), then

\[
G(w_1, \ldots, w_d) = \frac{2}{\sum_{i=1}^d w_i} \cdot w_d - \frac{d}{(\sum_{i=1}^d w_i)^2} \cdot w_d^2 - \frac{1}{\sum_{i=1}^d w_i} \cdot \frac{\sum_{i=1}^d w_i^{-\alpha}}{\sum_{i=1}^d w_i^{-(\alpha+1)}}.
\]

By direct calculation, we can check that \(G(w_1, \ldots, w_d)\) has zero gradient at \(W\), i.e. \(\nabla G|_W = 0\). Then we can calculate the Hessian matrix of \(G(w_1, \ldots, w_d)\) at \(W\)

\[
H|_W = \frac{2}{d^2 w^2} \begin{pmatrix}
\frac{\alpha+2}{d} + \alpha + 1 & \ldots & -\frac{\alpha+2}{d} \\
\vdots & \ddots & \vdots \\
-\frac{\alpha+2}{d} & \ldots & 1 - \frac{\alpha+2}{d}
\end{pmatrix},
\]

Set

\[
P = \begin{pmatrix}
\frac{\alpha+2}{d} + \alpha + 1 & \ldots & -\frac{\alpha+2}{d} \\
\vdots & \ddots & \vdots \\
-\frac{\alpha+2}{d} & \ldots & 1 - \frac{\alpha+2}{d}
\end{pmatrix},
\]

and

\[
Q = \begin{pmatrix}
\frac{\alpha+2}{d} & \ldots & -\frac{\alpha+2}{d} \\
\vdots & \ddots & \vdots \\
-\frac{\alpha+2}{d} & \ldots & 1 - \frac{\alpha+2}{d}
\end{pmatrix}.
\]

Notice that \(H|_W, P\) and \(Q\) satisfy

\[
H|_W = \frac{2}{d^2 w^2} P, \quad P = (\alpha + 1)I + Q,
\]

where \(I\) is the identity matrix. By direct calculation, we can get all the eigenvalues of matrix \(Q\)

\[
\lambda_1^Q = \ldots = \lambda_{d-2}^Q = 0, \lambda_{d-1}^Q = -\alpha + 1, \lambda_d^Q = -(d - 1),
\]

and then get all the eigenvalues of \(P\) by shifting \(Q\)’s eigenvalues by \(\alpha + 1\)

\[
\lambda_1^P = \ldots = \lambda_{d-2}^P = \alpha + 1, \lambda_{d-1}^P = 0, \lambda_d^P = \alpha + 2 - d.
\]

It is easy to see that when \(\alpha > d - 2\), one of \(P\)’s eigenvalues is zero and all the others are strictly positive. It is also easy to check that the sum of each row of \(P\) is zero, which means

\[
P \cdot (1, \ldots, 1)^T = 0 \cdot (1, \ldots, 1)^T.
\]
That is, the diagonal is an eigenvector associated with \( P \)'s zero eigenvalue and then \( H \)'s zero eigenvalue. Then we can conclude that \( G(w_1, \ldots, w_d) \) is convex along all the other directions except the diagonal, and hence the diagonal is its local minimum. \( \square \)

Keeping the notations of Lemma 5.2, we have the following lemma.

**Lemma 5.3.** For any fixed positive integer \( d \geq 3 \), there exists some \( \alpha_0(d) \), such that when \( \alpha > \alpha_0(d) \), \( U \) (uniform distribution) is a global minimum of \( g(u_1, \ldots, u_d) \) in \( \Delta \).

**Proof.** It is equivalent to show that for any \( u = (u_1, \ldots, u_d) \in \overset{\circ}{\Delta} \) the following holds

\[
2 \min_{i \in V} u_i - d \left( \min_{i \in V} u_i \right)^2 \geq \frac{\sum_{i=1}^{d} u_i^{\alpha}}{\sum_{i=1}^{d} u_i^{-(\alpha+1)}},
\]

(5.5)

with equality if and only if \( u \) is the uniform distribution, i.e. \( u = U \).

We will divide the proof of (5.5) into two cases:

1. \( u \) is in a neighborhood of uniform distribution;
2. \( u \) is bounded away from the uniform distribution. Equivalently, there exists some \( 0 < \kappa < 1 \) s.t. \( \min_{i \in V} u_i < \kappa/d \).

Case (1) directly follows from Lemma 5.2.

To prove case (2), first we try to use the minimum coordinates of \( u \) to bound the right hand side of (5.5) from above. More precisely, for fixed \( d \) and \( \alpha \), we will show that for any \( u \in \Delta \) the following holds

\[
\frac{\sum_{i=1}^{d} u_i^{\alpha}}{\sum_{i=1}^{d} u_i^{-(\alpha+1)}} \leq d^{1/(\alpha+1)} \min_{i \in V} u_i.
\]

(5.6)

Without loss of generality, we can assume \( u_d = \min_{i \in V} u_i \), and then if we set \( a_i = \min_{i \in V} u_i/u_i = u_d/u_i \in (0,1] \), (5.6) is equivalent to the following inequality with \( a_i \in (0,1] (i = 1, \ldots, d-1) \)

\[
1 + \frac{\sum_{i=1}^{d-1} a_i^\alpha}{\sum_{i=1}^{d-1} a_i^{1+\alpha}} \leq d^{1/(\alpha+1)}.
\]

(5.7)

To prove (5.7), observe that by Hölder’s inequality,

\[
\left( \frac{1 + \sum_{i=1}^{d-1} a_i^\alpha}{d} \right)^{1/\alpha} \leq \left( \frac{1 + \sum_{i=1}^{d-1} a_i^{\alpha+1}}{d} \right)^{1/(\alpha+1)},
\]

i.e.

\[
1 + \sum_{i=1}^{d-1} a_i^\alpha \leq d^{1/(\alpha+1)} \left( 1 + \sum_{i=1}^{d-1} a_i^{\alpha+1} \right)^{\alpha/(\alpha+1)}.
\]
Then
\[
1 + \frac{\sum_{i=1}^{d-1} a_i^{\alpha}}{1 + \sum_{i=1}^{d-1} a_i^{1+\alpha}} \leq \frac{d^{1/(\alpha+1)} \left(1 + \sum_{i=1}^{d-1} a_i^{\alpha+1}\right)^{\alpha/(\alpha+1)}}{1 + \sum_{i=1}^{d-1} a_i^{1+\alpha}} \leq \frac{d^{1/(\alpha+1)} \left(1 + \sum_{i=1}^{d-1} a_i^{1+\alpha}\right)^{1/(\alpha+1)}}{\left(1 + \sum_{i=1}^{d-1} a_i^{1+\alpha}\right)^{1/(\alpha+1)}} < d^{1/(\alpha+1)},
\]
proving (5.6). Notice that when \( \alpha \geq \log d/\log(2-\kappa) - 1 \), for any \( 0 < u_d < \kappa/d \) the following inequality holds
\[
d^{1/(\alpha+1)}, u_d < 2u_d - du_d^2, \quad \text{i.e.} \quad d^{1/(\alpha+1)} < 2 - du_d.
\]
Then (5.6) and (5.8) together imply that when \( \alpha \geq \log d/\log(2-\kappa) - 1 \), for any \( u \) satisfying \( \min_{i \in V} u_i < \kappa/d \)
\[
2 \min_{i \in V} u_i - d \left( \min_{i \in V} u_i \right)^2 > \frac{\sum_{i=1}^{d} u_i^{-\alpha}}{\sum_{i=1}^{d} u_i^{-(\alpha+1)}},
\]
(5.9)

Finally, we need to glue the results in the above two cases together. From case (1), for fixed \( d \) and \( \alpha > d - 2 \), there exists a neighborhood of the uniform distribution \( \mathcal{N}(U, \epsilon_0) \) s.t. for any \( u \in \mathcal{N}(U, \epsilon_0) \), (5.5) holds. For any \( u \neq U \) in \( \Delta \), since \( \sum_{i=1}^{d} u_i^{-\alpha} / \sum_{i=1}^{d} u_i^{-(\alpha+1)} \) is a decreasing function\(^2\) in \( \alpha, g(u_1, \ldots, u_d) \) in Lemma 5.2 is an increasing function in \( \alpha \). This allows us to take some common neighborhood \( \mathcal{N}(U, \epsilon_0) = \bigcap_{\alpha > d - 1} \mathcal{N}(U, \epsilon_0) \) just depending on \( d \) such that (5.5) holds. Take some \( \kappa = \kappa(d) < 1 \) such that
\[
\left\{ u \in \overset{\circ}{\Delta} : \min_{i \in V} u_i < \frac{\kappa}{d} \right\} \cup \mathcal{N}(U, \epsilon_0) = \overset{\circ}{\Delta}.
\]
Then set
\[
\alpha_0(d) = \max \left\{ d - 1, - \frac{\log d}{\log(2-\kappa)} - 1 \right\}.
\]
When \( \alpha > \alpha_0(d) \), the above two cases imply (5.5).

\[\begin{aligned}
\text{Lemma 5.4.} \quad & \text{For any fixed positive integer } d \geq 3, \text{ there exists some } \alpha_0(d), \text{ such that when } \alpha > \alpha_0(d), \text{ for any } (u, v) = (u_1, \ldots, u_d, v_1, \ldots, v_d) \in \bar{D} \text{ the following holds}
\quad & 2 \sum_{i=1}^{d} u_i v_i \geq \frac{\sum_{i=1}^{d} u_i^{-\alpha}}{\sum_{i=1}^{d} u_i^{-(\alpha+1)}} + \frac{\sum_{i=1}^{d} v_i^{-\alpha}}{\sum_{i=1}^{d} v_i^{-(\alpha+1)}}, \\
& \text{with equality if and only if } u_i = v_i = \frac{1}{2} \text{ for any } i, \text{ i.e. } (u, v) = (U, U).
\end{aligned}\]

\[\begin{aligned}
\text{Proof.} \quad & \text{First for any fixed } u, v \in \overset{\circ}{\Delta}, \text{ we will bound the left hand side of (5.10) from below by the minimum coordinates of } u \text{ and } v. \text{ More precisely, if we construct two } d-\text{dimensional vectors } u' \text{ and } v' \text{ by the minimum coordinates of } u \text{ and } v
\quad & u' = \left( \min_{i \in V} u_i, \ldots, \min_{i \in V} u_i, 1 - (d - 1) \min_{i \in V} u_i \right)
\end{aligned}\]

\[\text{This can be proved by directly checking the derivative.}\]
and
\[ v' = \left( 1 - (d - 1) \min_{i \in V} v_i, \min_{i \in V} v_i, \ldots, \min_{i \in V} v_i \right), \]
we will show
\[ \sum_{i=1}^d u_i v_i \geq \sum_{i=1}^d u'_i v'_i = \min_{i \in V} u_i + \min_{i \in V} v_i - d \cdot \min_{i \in V} u_i \min_{i \in V} v_i. \quad (5.11) \]

By the Rearrangement inequality, it suffices to show (5.11) for any \((u, v) \in \partial \hat{\Delta}\) satisfying \(u_1 \leq \ldots \leq u_d\) and \(v_1 \geq \ldots \geq v_d\). For such \(u\) and \(v\), we have
\[
\sum_{i=1}^d u_i v_i = u_1 v_1 + \sum_{i=2}^d u_i v_i + \sum_{i=2}^d u_i (v_i - v_d) \geq u_1 v_1 + \sum_{i=2}^d u_i v_i + \sum_{i=2}^d u_1 (v_i - v_d) = u_1 (1 - (d - 1)v_d) + \sum_{i=2}^d u_i v_d = \sum_{i=1}^d u_i v'_i \geq \sum_{i=1}^d u'_i v'_i,
\]
where the last step is obtained by repeating the same argument as in the previous steps. Then we have proved (5.11).

By Lemma 5.3 we know that there exists some \(\alpha_0(d)\), such that when \(\alpha > \alpha_0(d)\), for any \((u, v) \in \partial \hat{\Delta}\) the following inequalities hold
\[
2 \min_{i \in V} u_i - d \left( \min_{i \in V} u_i \right)^2 \geq \frac{\sum_{i=1}^d u_i - \alpha}{\sum_{i=1}^d u_i - (\alpha + 1)}
\]
and
\[
2 \min_{i \in V} v_i - d \left( \min_{i \in V} v_i \right)^2 \geq \frac{\sum_{i=1}^d v_i - \alpha}{\sum_{i=1}^d v_i - (\alpha + 1)}.
\]
Then by the following
\[
2 \left( \min_{i \in V} u_i + \min_{i \in V} v_i - d \min_{i \in V} u_i \min_{i \in V} v_i \right) \geq \left( 2 \min_{i \in V} u_i - d \left( \min_{i \in V} u_i \right)^2 \right) + \left( 2 \min_{i \in V} v_i - d \left( \min_{i \in V} v_i \right)^2 \right),
\]
we have proved (5.10).

It is also easy to check the equality holds if and only if both \(u\) and \(v\) are uniform distributions, i.e. \(u = v = U\). \(\square\)

**Proof of Lemma 5.1** Recall that \(f(x) = \delta 1_{x \leq \delta} + x 1_{x > \delta}\). The proof of this lemma is divided into three cases:

1. \(\min_{i \in V} f(u_i) > \delta\) and \(\min_{i \in V} f(v_i) > \delta\);
2. \(\min_{i \in V} f(u_i) = \delta\) and \(\min_{i \in V} f(v_i) > 2\delta\) (or the symmetric case);
3. \(\min_{i \in V} f(u_i) = \delta\) and \(\min_{i \in V} f(v_i) \leq 2\delta\) (or the symmetric case).
To prove case (1), observe that \(\min_{i \in V} f(u_i) > \delta\) and \(\min_{i \in V} f(v_i) > \delta\) imply \(f(u_i) = u_i > \delta\) and \(f(v_i) = v_i > \delta\) for any \(i \in V\). Hence, to prove (5.3), it is equivalent to prove (5.10). Then by Lemma 5.4, when \(\alpha > \alpha_0(d) + 1\), \(\frac{d}{dt}(L(\Phi_t(x)))\big|_{(u,v)} \leq 0\), with equality if and only if \((u, v) = (U, U)\).

Let us prove case (2). From (5.2), it follows that

\[
\frac{d}{dt}(L(\Phi_t(x)))\big|_{(u,v)} = -2 \sum_{i=1}^{d} u_i v_i + \sum_{i=1}^{d} \frac{u_i f(u_i)^{-\alpha}}{\sum_{k=1}^{d} f(u_k)^{-\alpha}} + \sum_{i=1}^{d} \frac{v_i f(v_i)^{-\alpha}}{\sum_{k=1}^{d} f(v_k)^{-\alpha}}
\]

\[
\leq -2 \sum_{i=1}^{d} u_i v_i + \sum_{i=1}^{d} \frac{f(u_i) f(u_i)^{-\alpha}}{\sum_{k=1}^{d} f(u_k)^{-\alpha}} + \sum_{i=1}^{d} \frac{f(v_i) f(v_i)^{-\alpha}}{\sum_{k=1}^{d} f(v_k)^{-\alpha}}
\]

\[
\leq -2 \sum_{i=1}^{d} u_i v_i + d^{1/\alpha} \left( \min_{i \in V} f(u_i) + \min_{i \in V} f(v_i) \right),
\]

(5.12)

where the last step is by (5.3), which actually holds for any collection of positive numbers. Since

\[
\sum_{i=1}^{d} u_i v_i \geq \max \left\{ \min_{i \in V} u_i, \min_{j \in V} v_j \right\},
\]

it follows from the conditions in case (2) that

\[
\sum_{i=1}^{d} u_i v_i \geq \max \left\{ \min_{i \in V} u_i, \min_{j \in V} v_j \right\} \geq \min_{j \in V} v_j = \min_{i \in V} f(v_i).
\]

Hence,

\[
\frac{d}{dt}(L(\Phi_t(x)))\big|_{(u,v)} \leq -2 \sum_{i=1}^{d} u_i v_i + d^{1/\alpha} \left( \min_{i \in V} f(u_i) + \min_{i \in V} f(v_i) \right)
\]

\[
\leq -2 \min_{i \in V} f(v_i) + d^{1/\alpha} \left( \delta + \min_{i \in V} f(v_i) \right)
\]

\[
= - \left( 2 - d^{1/\alpha} \right) \min_{i \in V} f(v_i) + \frac{d}{2} \delta 
\]

\[
\leq - \left( 2 - d^{1/\alpha} \right) 2\delta + \frac{d^{1/\alpha}}{2} \delta = - \left( 4 - 3d^{1/\alpha} \right) \delta.
\]

So one can choose \(\alpha > \log d/\log 4\) such that \(4 - 3d^{1/\alpha} > 0\), then obtaining \(\frac{d}{dt}(L(\Phi_t(x)))\big|_{(u,v)} < 0\).

To prove case (3), one can choose \(\alpha > \log_2 d\) such that \(2d^{1/\alpha} < 3\). Then by the definition of \(D^\delta\), one has

\[
L(u, v) \geq 3\delta > \frac{3}{2} d^{1/\alpha} \delta, \quad \forall (u, v) \in D^\delta.
\]

(5.13)

(5.12) and (5.13) imply

\[
\frac{d}{dt}(L(\Phi_t(x)))\big|_{(u,v)} \leq -2 \sum_{i=1}^{d} u_i v_i + 3d^{1/\alpha} \delta < 0.
\]

(5.14)
To sum up, taking
\[ \alpha(d) = \max \left\{ \alpha_0(d) + 1, \log d / \log \frac{4}{3}, \log_2 d \right\}, \]
we establish the lemma. \( \square \)

5.1.2. The main lemma. Now it comes to our main lemma to characterize the chain recurrent set for our specific semiflow \( \Phi \).

Lemma 5.5. Assume the result in Lemma 5.4. Let
\[ S^\delta = \{(u, v) \in D : \sum_{i=1}^{d} u_i v_i \leq 4\delta \}. \]
Then \( \text{CR}(\Phi) \subset S^\delta \cup (U, U) \).

Proof. Set \( \zeta_0 = 3.5\delta, \zeta_1 = 4\delta \) and \( \zeta_2 = 1 \). Define
\[ M_j = \{(x, y) \in D : L(x, y) \leq \zeta_j \}, \quad j = 0, 1, 2. \]
Note that \( M_1 = S^\delta, M_2 = D \). By Lemma 5.1 and Proposition 4.8, \( M_j(j = 0, 1, 2) \) are compact invariant sets. Clearly, the lemma will follow once we prove:
(a) \( CR_1 = \text{CR}(\Phi) \cap M_1 \) and \( CR_2 = \text{CR}(\Phi) \cap (M_2 \setminus M_1) \) are invariants sets;
(b) \( CR_2 = (U, U) \).

Let’s prove (a).
By the invariance of \( \text{CR}(\Phi) \) and \( M_1 \), it is clear that \( CR_1 \) is invariant.

We will show \( CR_2 \) is invariant by contradiction. Suppose for some \( z \in CR_2 \), there exists some \( T_0 > 0 \), s.t. \( \Phi_{T_0}(z) \in M_1 \). Then by Lemma 5.1 and compactness of \( M_1 \setminus M_0 \), there exists some \( T_1 > T_0 \), such that
\[ L(\Phi_{T_1}(z)) < \zeta_0, \text{ i.e. } \Phi_{T_1}(z) \in M_0. \] (5.15)
Also by Lemma 5.1 and compactness of \( M_1 \setminus M_0 \), there exists some \( T_2 > 0 \), s.t.
\[ \Phi_{T_2}(M_1) \subset M_0. \]

Let \( \rho_0 = \text{dist}(M_0, D \setminus M_1) > 0 \) and \( T = \max\{T_1, T_2\} \). By assumption, there are points \( z_0 = z, z_1, \ldots, z_{k-1}, z_k = z \in D \) and real numbers \( t_0, \ldots, t_{k-1} > T \) such that
\[ \text{dist}(\Phi_{t_i}(z_i), z_{i+1}) < \rho_0, \quad i = 0, \ldots, k-1. \] (5.16)
Thus \( \Phi_{t_0}(z_0) = \Phi_{t_0}(z) \in M_0 \) and so, by (5.16), \( z_1 \in M_1 \). By induction, we claim that \( z_1, z_2, \ldots, z_k \in M_1 \). Indeed, if \( z_i \in M_1 \), \( \Phi_{t_i}(z_i) \in M_0 \), and then by (5.16), \( z_{i+1} \in M_1 \). In particular, \( z_k = z \in M_1 \), which contradicts. Hence, \( \Phi_t(CR_2) \subset M_2 \setminus M_1 \) for all \( t \geq 0 \). By invariance of \( \text{CR}(\Phi) \), \( CR_2 \) is invariant.

It remains to prove (b).
Since \( (U, U) \in CR_2 \), it suffices to show \( CR_2 \subset (U, U) \). For any \( z \in CR_2 \), by invariance of \( CR_2 \) and the non-increasing property of the Lyapunov function \( L(\cdot) \) along any trajectory in \( M_2 \setminus M_1 \), it follows that the limit of \( L(\Phi_t(z)) \) exists. Set
\[ L(\Phi_\infty(z)) = \lim_{t \to +\infty} L(\Phi_t(z)). \]
Then for any \( p \in \omega(z) \) (the omega limit set of \( z \)), \( L(p) = L(\Phi_\infty(z)) \). Together with invariance of \( \omega(z) \), this implies \( L(\cdot) \) is constant along trajectories in \( \omega(z) \). Therefore, \( \omega(z) \subset (U, U) \). Since \( \omega(z) \) is nonempty, \( \omega(z) = (U, U) \).
Now we will show that \( z = (U, U) \) by contradiction. Suppose \( z \neq (U, U) \), then there exists a neighborhood of \( z \), s.t. \( L(\Phi_t(z)) \) is strictly decreasing in this neighborhood. Since

\[
L(\Phi_t(z)) \geq L(\omega(z)) = L(U, U) = 1/d,
\]

there exists some \( \epsilon > 0 \), such that \( L(z) > 1/d + \epsilon \). Define

\[
E = \left\{ w \in D \mid L(w) \leq \frac{1}{d} + \frac{\epsilon}{3} \right\}, \quad F = \left\{ w \in D \mid L(w) \geq \frac{1}{d} + \frac{\epsilon}{2} \right\}.
\]

By Lemma 5.1 and compactness of \( F \), there exists some \( T \), such that \( \Phi_T(F) \subset E \). Let \( \rho_1 = \text{dist}(E, F) > 0 \). Then by a similar argument as that in (a), one can get a contradiction. \( \square \)

5.2. Non-convergence to unstable equilibrium. By Theorem 4.7 and Lemma 5.6,

\[
\mathbb{P} \left( L(\{z(n)\}_{n \geq 0}) \subset S^\delta \cup (U, U) \right) = 1.
\]

Since \( S^\delta \) and \((U, U)\) are disconnected, we can finish the proof of Theorem 2.2 by proving the following lemma.

**Lemma 5.6.** When \( \alpha > 1 \), for any initial condition \((x(0), y(0)) \in D, z(n) \) in (2.8) satisfies

\[
\mathbb{P} \left( \lim_{n \to \infty} z(n) = (U, U) \right) = 0.
\]

**Lemma 5.6** is an application of the following theorem to our case. This theorem is due to Pemantle [15] which, to our purpose, is stated as

**Theorem 5.7.** [15, Theorem 1] Define a stochastic process

\[
z(n + 1) - z(n) = \frac{1}{n + 1 + d} [F(z(n)) + u_n]
\]

with \( \mathbb{E}(u_n | F_n) = 0 \) and such that \( z(n) \) always remains in bounded domain \( D \). Let \( p \) be any point in \( D \) with \( F(p) = 0 \), let \( \mathcal{N} \) be a neighborhood of \( p \) and assume that there are constants \( c_1, c_2 > 0 \) for which the following conditions are satisfied whenever \( z(n) \in \mathcal{N} \) and \( n \) is sufficiently large:

(1) \( p \) is a linearly unstable critical point,

(2) \( \mathbb{E}(u_n \cdot \theta | F_n) \geq c_1 \) for every unit vector \( \theta \in TD \) (see the definition of \( TD \) in Notation 2.7),

(3) \( \|u_n\| \leq c_2 \),

where \((u_n \cdot \theta)^+ = \max\{u_n \cdot \theta, 0\} \) is the positive part of \( u_n \cdot \theta \). Assume \( F \) is smooth enough to apply the stable manifold theorem: at least \( C^2 \). Then

\[
\mathbb{P} \left( \lim_{n \to \infty} z(n) = p \right) = 0.
\]

The rest of this section is to verify that our \( z(n) \) in (2.8) falls into the setting of Theorem 5.7 with \( p = (U, U) \). First it is easy to check that \( p \) is a critical point of the vector field \( F \) in (4.3), i.e. \( F(U, U) = 0 \). After introducing the following definition, we will show that \( p \) is a linearly unstable critical point.

**Definition 5.8.** Let \( T \) be the linear approximation to some vector field \( F \) near a critical point \( p \) so that \( F(p + w) = T(w) + O(|w|^2) \), then

(a) If all the eigenvalues of \( T \) have strictly negative real part, \( p \) is called an attracting point.
(b) If some eigenvalues of $T$ have strictly positive real part, $p$ is called an unstable critical point.

**Lemma 5.9.**

1. When $\alpha > 1$, $(U,U)$ is a linearly unstable critical point for $F$ in (4.3).
2. When $\alpha < 1$, $(U,U)$ is an attracting point for $F$ in (4.3).

**Proof.** By Taylor expansion of $F$ in a neighborhood of $(U,U)$, we have

$$F(p + w) = DF|_p \cdot w + O(|w|^2),$$

where $DF$ is Jacobian matrix and $w$ is some vector in a neighborhood of 0 (2d dimensional vector). By direct calculation, one can have

$$DF|_p = \begin{pmatrix}
-1 & 0 & \ldots & 0 \\
0 & -1 & 0 & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & -1 \\
-\alpha + \frac{\alpha}{d} & \frac{\alpha}{d} & \ldots & \frac{\alpha}{d} \\
\frac{\alpha}{d} & -\alpha + \frac{\alpha}{d} & \frac{\alpha}{d} & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\alpha}{d} & \ldots & \frac{\alpha}{d} & -\alpha + \frac{\alpha}{d}
\end{pmatrix}.$$

In order to get all the eigenvalues of $DF|_p$, one need to solve

$$|DF|_p - \lambda I_{2d \times 2d} = 0,$$

where $|\cdot|$ represents the matrix’s determinant and $I_{2d \times 2d}$ is a 2d dimensional identity matrix. Notice that $DF|_p$ has the same upper-right and lower-left block matrix. We set this $d \times d$ matrix as

$$B = \begin{pmatrix}
-\alpha + \frac{\alpha}{d} & \frac{\alpha}{d} & \ldots & \frac{\alpha}{d} \\
\frac{\alpha}{d} & -\alpha + \frac{\alpha}{d} & \frac{\alpha}{d} & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\alpha}{d} & \ldots & \frac{\alpha}{d} & -\alpha + \frac{\alpha}{d}
\end{pmatrix}.$$

Because the sum of $B$’s each row is zero, $|B| = 0$. Then one can easily check $\lambda = -1$ is a solution to (5.17). Now assume $\lambda \neq -1$. By Schur complement, from (5.17) we get

$$|B^2 - (\lambda + 1)^2 I_{d \times d}| = 0,$$

where $I_{d \times d}$ is a $d$ dimensional identity matrix. (5.18) is equivalent to

$$|B - (\lambda + 1)I_{d \times d}| \cdot |B + (\lambda + 1)I_{d \times d}| = 0.$$  

(5.19)

Under the assumption $\lambda \neq -1$, we easily get all the solutions to (5.19): $\lambda = -1 \pm \alpha$. Hence all the eigenvalues of $DF|_p$ without counting multiplicities are $-1, -1 \pm \alpha$.

So if $\alpha > 1$, $DF|_p$ has the positive eigenvalue $-1 + \alpha$; If $\alpha < 1$, all of its eigenvalues are strictly negative. This concludes the proof. \hfill \square

Clearly, $u_n$ in (4.2) satisfies condition (3) of Theorem 5.7. It remains to check condition (2), which is the following lemma.
Lemma 5.10. In a small neighborhood of $p = (U,U)$, there exists some constant $c > 0$, s.t. $\mathbb{E}((u_n \cdot \theta)^+|\mathcal{F}_n) \geq c$ for every unit vector $\theta = (\theta_k)_{1 \leq k \leq 2d} \in TD$.

Proof. Conditioning on $X_{n+1} = i \in V, Y_{n+1} = j \in V$, one has
\begin{align*}
u_n \cdot \theta &= (1 - \pi_1(y(n)))\theta_i - \sum_{m \neq i, m \in V} \pi_m(y(n))\theta_m \\
&\quad + (1 - \pi_j(x(n)))\theta_{j+d} - \sum_{s \neq j, s \in V} \pi_s(x(n))\theta_{s+d}
&= \theta_i + \theta_{j+d} - \sum_{m \in V} \pi_m(y(n))\theta_m - \sum_{s \in V} \pi_s(x(n))\theta_{s+d}. \quad (5.20)
\end{align*}

Now we will show for any unit vector $\theta \in TD$, the following holds
\begin{equation}
\max_{1 \leq k \leq 2d} \theta_k \geq \frac{1}{2d(d-1)}. \quad (5.21)
\end{equation}

Observe that $\theta$ as a unit vector ($\|\theta\| = 1$) always satisfies $\max_{1 \leq k \leq 2d} |\theta_k| \geq \frac{1}{2d}$. Then if the following inequality doesn’t hold
\begin{equation}
\max_{1 \leq k \leq 2d} \theta_k \geq \frac{1}{2d} \left( > \frac{1}{2d(d-1)} \right),
\end{equation}

there must be $\min_{1 \leq k \leq 2d} \theta_k \leq -1/(2d)$. Then it follows from $\sum_{k=1}^{d} \theta_k = 0$ and $\sum_{k=d+1}^{2d} \theta_k = 0$ that there exists some coordinate $1 \leq k_0 \leq 2d$, s.t. $\theta_{k_0} \geq 1/[2d(d-1)]$, which implies (5.21). From (5.21), without loss of generality, we can assume
\begin{equation}
\theta_1 = \max_{1 \leq k \leq 2d} \theta_k \geq \frac{1}{2d(d-1)}. \quad (5.22)
\end{equation}

Then by $\sum_{k=d+1}^{2d} \theta_k = 0$, there also exists some $j_0 \in V$, s.t. $\theta_{j_0+d} \geq 0$.

Because $(x(n), y(n))$ lives in a small neighborhood of $(U,U)$, $\pi(x(n))$ and $\pi(y(n))$ also live in a small neighborhood of $(U,U)$, and hence in (5.21),
\begin{equation}
\sum_{m \in V} \pi_m(y(n))\theta_m \sim 0, \quad \sum_{s \in V} \pi_s(x(n))\theta_{s+d} \sim 0.
\end{equation}

Therefore, by (5.20),
\begin{align*}
\mathbb{E}((u_n \cdot \theta)^+|\mathcal{F}_n) &\geq \mathbb{P}(X_{n+1} = 1, Y_{n+1} = j_0|\mathcal{F}_n)\theta_1 \\
&= \mathbb{P}(X_{n+1} = 1|\mathcal{F}_n)\mathbb{P}(Y_{n+1} = j_0|\mathcal{F}_n)\theta_1.
\end{align*}

Then by the same fact that $\pi(x(n))$ and $\pi(y(n))$ live in a small neighborhood of $(U,U)$, both $\mathbb{P}(X_{n+1} = 1|\mathcal{F}_n)$ and $\mathbb{P}(Y_{n+1} = j_0|\mathcal{F}_n)$ are close to $1/d$. Then together with (5.22), it follows that $\mathbb{E}((u_n \cdot \theta)^+|\mathcal{F}_n)$ is uniformly bounded from below by some positive constant. This completes the proof. \hfill \Box

Now we can apply Theorem 5.7 obtaining Lemma 5.6.

6. Further problems

This paper is part of a project to answer the following

Problem 6.1. For the model defined in (2.1) and (2.3), what is the random walks’ behavior when choosing different graphs and weights?
One concrete version of the above problem can be the nearest-neighbor interacting random walks on \( \mathbb{Z}^d \) with the weight sequence \( w_k = k^{-\alpha} \) (\( \alpha > 0 \)). Problem 6.1 remains widely open in general.

Now we make a conjecture regarding the model in (2.6) and (2.7) we have studied.

**Conjecture 6.2.** For any positive integer \( d \geq 3 \), \( \alpha_c = 1 \) is critical, i.e. for any given initial condition \( (x(0), y(0)) \in D \),

1. when \( \alpha > 1 \), there exists some constant \( c = c(\alpha, d) \) (not large) depending on \( \alpha \) and \( d \), such that the following holds

\[
P \left\{ \exists n_0, \bigcap_{n \geq n_0} \left\{ \sum_{i=1}^{d} x_i(n)y_i(n) \leq c\delta \right\} \right\} = 1.
\]

2. when \( 0 < \alpha < 1 \), the following holds

\[
P \left\{ \lim_{n \to \infty} z(n) = (U, U) \right\} = 1.
\]

Another problem of interest is

**Problem 6.3.** When \( \delta = 0 \) in (2.6) and (2.7), how to derive the similar result as Theorem 2.2?

7. **Acknowledgements**

This work was supported by the ISF. The author thanks Itai Benjamini for his very stimulating questions and Gady Kozma for reading the draft and giving many comments. The author is indebted to Ofer Zeitouni for his patient guidance and numerous insights on this work. The author is also thankful to Omri Sarig and Yuri Lima for all the discussions on dynamical systems.

**References**

[1] O. Angel, N. Crawford, and G. Kozma, *Localization for linearly edge reinforced random walks*, arXiv preprint arXiv:1203.4010 (2012).

[2] M. Benaïm, *A dynamical system approach to stochastic approximations*, SIAM J. Control Optim. 34 (1996), no. 2, 437–472.

[3] M. Benaïm, *Dynamics of stochastic approximation algorithms*, Séminaire de Probabilités, XXXIII (1999), 1–68.

[4] A. Benveniste, P. Priouret, and M. Métivier, *Adaptive algorithms and stochastic approximations*, Applications of Mathematics (New York), Springer-Verlag, Berlin, 1990.

[5] B. Davis, *Reinforced random walk*, Probab. Theory Related Fields 84 (1990), no. 2, 203–229.

[6] R. Durrett, *Probability: theory and examples*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2010.

[7] B.M. Hill, D. Lane, and W. Sudderth, *A strong law for some generalized urn processes*, Ann. Probab. 8 (1980), no. 2, 214–226.

[8] J. Kiefer and J. Wolfowitz, *Stochastic estimation of the maximum of a regression function*, Ann. Math. Statistics 23 (1952), no. 3, 462–466.

[9] Y. Kovchegov, *Multi-particle processes with reinforcements*, J. Theoret. Probab. 21 (2008), no. 2, 437–448.

[10] H.J. Kushner and D.S. Clark, *Stochastic approximation methods for constrained and unconstrained systems*, Applied Mathematical Sciences, vol. 6, Springer-Verlag, New York-Berlin, 1978.

[11] V. Limic, *Attracting edge property for a class of reinforced random walks*, Ann. Probab. 31 (2003), no. 3, 1615–1654.

[12] V. Limic and P. Tarrès, *Attracting edge and strongly edge reinforced walks*, Ann. Probab. 35 (2007), no. 5, 1783–1806.
[13] L. Ljung, *Analysis of recursive stochastic algorithms*, IEEE Trans. Automatic Control **22** (1977), no. 4, 551–575.

[14] R. Pemantle, *Phase transition in reinforced random walk and RWRE on trees*, Ann. Probab. **16** (1988), no. 3, 1229–1241.

[15] __L. Ljung__, *Nonconvergence to unstable points in urn models and stochastic approximations*, Ann. Probab. **18** (1990), no. 2, 698–712.

[16] __L. Ljung__, *A survey of random processes with reinforcement*, Probab. Surv. **4** (2007), no. 1-79.

[17] R. Pemantle and S. Volkov, *Vertex-reinforced random walk on Z has finite range*, Ann. Probab. **27** (1999), no. 3, 1368–1388.

[18] H. Robbins and S. Monro, *A stochastic approximation method*, Ann. Math. Statistics **22** (1951), 400–407.

[19] C. Sabot and P. Tarrès, *Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model*, arXiv preprint arXiv:1111.3991 (2011).

[20] T. Sellke, *Reinforced random walk on the d-dimensional integer lattice*, preprint (1994).

[21] P. Tarrès, *Vertex-reinforced random walk on Z eventually gets stuck on five points*, Ann. Probab. **32** (2004), no. 3B, 2650–2701.

[22] S. Volkov, *Phase transition in vertex-reinforced random walks on with non-linear reinforcement*, J. Theoret. Probab. **19** (2006), no. 3, 691–700.

**Weizmann Institute of Science, Faculty of Mathematics and Computer Science, POB 26, 76100, Rehovot, Israel.**

*E-mail address: jun.chen@weizmann.ac.il*