Electronic Journal of Probability

Electron. J. Probab. **19** (2014), no. 113, 1–17.
ISSN: 1083-6489 DOI: 10.1214/EJP.v19-2669

Two particles' repelling random walks on the complete graph

Jun Chen*

Abstract

We consider two particles' repelling random walks on complete graphs. In this model, each particle has higher probability to visit the vertices which have been seldom visited by the other one. By a dynamical approach we prove that the two particles' occupation measure asymptotically has small joint support almost surely if the repulsion is strong enough.

Keywords: Repelling random walks; Reinforced random walk; multi-particle; complete graph; stochastic approximation algorithms; dynamical approach; chain recurrent set; Lyapunov function.
AMS MSC 2010: Primary 60K35, Secondary 37C10.
Submitted to EJP on March 10, 2013, final version accepted on December 2, 2014.

1 Introduction and statement of result

In this paper, we consider a model of multi-particle vertex *repelling random walks*, which is analogous to the well-studied *reinforced random walks* (RRW). See [13] for a general reference to RRW. To the best knowledge of the author, this paper is one of the first papers [7] investigating multi-particle interacting random walks. Our model was proposed by Itai Benjamini around the year 2010 and can be generalized to any graph.

Now we define the model of two particles' repelling random walks on a complete graph. Denote the two particles by X and Y, and let $G = (V, E)$ be a complete graph with $V = \{1, \dots, d\}$. 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. Assume that $N(X, v, 0) = N(Y, v, 0) = 1$ for any $v \in V$. Let

$$x_i(n) = \frac{N(X,i,n)}{n+d},\ y_i(n) = \frac{N(Y,i,n)}{n+d},\ \forall i \in V \tag{1.1}$$

be X and Y's *empirical occupation measure* on V by time n. Let $\mathcal{F}_n (n \in \mathbb{N})$ be the natural filtration generated by $\{X_k, 0 \le k \le n\}$ and $\{Y_k, 0 \le k \le n\}$. Then we define the random walks (X_n, Y_n) by

$$\mathbb{P}(X_{n+1} = i | \mathcal{F}_n) = \frac{\left[\delta 1_{y_i(n) \le \delta} + y_i(n) 1_{y_i(n) > \delta}\right]^{-\alpha}}{\sum_{k=1}^{d} \left[\delta 1_{y_k(n) \le \delta} + y_k(n) 1_{y_k(n) > \delta}\right]^{-\alpha}}, \quad \forall i \in V, \tag{1.2}$$

*Division of the Humanities and Social Sciences, California Institute of Technology, USA.
E-mail: jchen3@caltech.edu

and

$$\mathbb{P}(Y_{n+1}=j|\mathcal{F}_n)=\frac{\left[\delta 1_{x_j(n)\le\delta}+x_j(n)1_{x_j(n)>\delta}\right]^{-\alpha}}{\sum_{k=1}^d\left[\delta 1_{x_k(n)\le\delta}+x_k(n)1_{x_k(n)>\delta}\right]^{-\alpha}},\quad \forall j\in V, \tag{1.3}$$

where δ,α are some fixed positive numbers and $1_{\{\cdot\}}$ is the indicator function. Notice that by the definition of (1.2) and (1.3), the random walks are lazy random walks, i.e. the particles could stay at their current locations.

Let $x(n)=(x_1(n),\dots,x_d(n))$, $y(n)=(y_1(n),\dots,y_d(n))$, and $z(n)$ be a $2d$ dimensional vector

$$z(n)=(x(n),y(n))=(x_1(n),\dots,x_d(n),y_1(n),\dots,y_d(n)). \tag{1.4}$$

Notice that $z(n)$ is a Markov chain living in $\mathbb{R}_+^{2d}$. We are interested in $z(n)$'s asymptotic behavior.

Here we want to mention that when $\min_{i,j\in V}\{x_i(n),y_j(n)\}>\delta$, (1.2) and (1.3) are equivalent to the following formulas

$$\mathbb{P}(X_{n+1}=i|\mathcal{F}_n)=\frac{N(Y,i,n)^{-\alpha}}{\sum_{k=1}^d N(Y,k,n)^{-\alpha}},\quad \forall i\in V, \tag{1.5}$$

$$\mathbb{P}(Y_{n+1}=j|\mathcal{F}_n)=\frac{N(X,j,n)^{-\alpha}}{\sum_{k=1}^d N(X,k,n)^{-\alpha}},\quad \forall j\in V, \tag{1.6}$$

which can be viewed as a multi-particle analogue of the classical RRW with nonlinear reinforcement. In the definition of (1.2) and (1.3), we are not able to work with $\delta=0$ due to a technical difficulty of our proof. See Problem 4.2.

Then we can state our main result.

Theorem 1.1. *For any fixed positive integer $d\ge 3$, there exists some $\alpha(d)$, s.t. when $\alpha\ge\alpha(d)$, for any fixed $\delta>0$ in the definition of (1.2) and (1.3), the two components $x(n)$ and $y(n)$ of $z(n)$ in (1.4) asymptotically have joint support bounded by 4δ almost surely, i.e.*

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

Our result says that for a fixed complete graph, when the repulsion is strong enough, in the definition of our model we can push δ down to zero to relax its restriction on the occupation measures $x(n)$ and $y(n)$, so that the joint support of the particles' occupation measures can be made arbitrarily small. Our result is analogous to the *localization* results [1, 3, 9, 10, 14, 16] in the RRW models.

The organization of this paper is as follows: In Section 2, we will do some preparation work for the proof of Theorem 1.1. More specifically, we will introduce a notion of *stochastic approximation algorithm*, describe the *dynamical approach* and then apply them to $z(n)$, finally 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 3, we will prove Theorem 1.1. In Section 4, we will propose some open problems.

2 Some preparations to prove the main result

2.1 Stochastic approximation algorithm and dynamical approach

A *stochastic approximation algorithm* is a discrete time stochastic process whose form can be written as

$$z(n+1)-z(n)=\gamma_n H(z(n),\xi(n)) \tag{2.1}$$

where $H:\mathbb{R}^m\times\mathbb{R}^k\to\mathbb{R}^m$ is a measurable function that characterizes the algorithm, $\{z(n)\}_{n\ge0}\subset\mathbb{R}^m$ is the sequence of parameters to be recursively updated, $\{\xi(n)\}_{n\ge0}\subset$

$\mathbb{R}^k$ is a sequence of random variables defined on some probability space, 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 [15] and Kiefer and Wolfowitz [6].

Observe that $z(n)$ in (1.4) is a stochastic approximation algorithm. Indeed, from (1.1)

$$\begin{aligned} x_i(n+1) - x_i(n) &= \frac{N(X,i,n) + 1_{X_{n+1}=i}}{n+1+d} - \frac{N(X,i,n)}{n+d} \\ &= \frac{-x_i(n) + 1_{X_{n+1}=i}}{n+1+d}. \end{aligned} \tag{2.2}$$

Similarly, a difference equation for $y_i(n)$ can be derived. Then $z(n)$ satisfies (2.1) with

$$\gamma_n = \frac{1}{n+1+d},\ \xi(n) = (1_{X_{n+1}=1}, \ldots, 1_{X_{n+1}=d}, 1_{Y_{n+1}=1}, \ldots, 1_{Y_{n+1}=d}) \tag{2.3}$$

and $H : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ be $H(z(n), \xi(n)) = -z(n) + \xi(n)$. That is,

$$z(n+1) - z(n) = \frac{1}{n+1+d}\left(-z(n) + \xi(n)\right). \tag{2.4}$$

The dynamical approach is a method used to analyze stochastic approximation algorithms, introduced by Ljung [11] and Kushner and Clark [8]. 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 [8], numerous articles by Kushner, and more recently the book by Benveniste, Metivier, and Priouret [4].

In the above perspective, our stochastic approximation algorithm $z(n)$ can be written as

$$z(n+1) - z(n) = \gamma_n \left\{(-z(n) + \mathbb{E}\left[\xi(n)|\mathcal{F}_n\right]) + (\xi(n) - \mathbb{E}\left[\xi(n)|\mathcal{F}_n\right])\right\}. \tag{2.5}$$

Before moving on, we need to introduce some notations.

Notation 2.1. *1. Let Δ be the closed $(d-1)-$dimensional simplex*

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

Denote the relative interior of Δ by $\overset{\circ}{\Delta}$.

2. *Let D be the product of two simplices $\Delta \times \Delta$*

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

Denote the relative interior of D by $\overset{\circ}{D}$, and the boundary of D by ∂D ;

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

$$TD = T(\Delta \times \Delta) = \left\{(u,v) \in \mathbb{R}^{2d},\ \sum_{i=1}^d u_i = 0,\ \sum_{i=1}^d v_i = 0\right\}.$$

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

5. *Let $\|\cdot\|$ be the L^1 norm on $\mathbb{R}^{2d}$.*

Let $\varpi(x) = [\delta 1_{x\le\delta} + x1_{x>\delta}]^{-\alpha}$, and define a map $\pi = (\pi_1, \ldots, \pi_d) : \Delta \to \Delta$ by

$$\pi_i(x) = \frac{\varpi(x_i)}{\sum_{k=1}^d \varpi(x_k)}, \quad \forall x \in \Delta. \tag{2.6}$$

Observe that, by (1.2), (1.3) and (2.3),

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

Thus, defining $\{u_n\}_{n\ge0} \subset \mathbb{R}^{2d}$ by

$$u_n = \xi(n) - \mathbb{E}\left[\xi(n)|\mathcal{F}_n\right] \tag{2.7}$$

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

$$F_i(x_1, \ldots, x_d, y_1, \ldots, y_d) = \begin{cases} -x_i + \pi_i(y_1, \ldots, y_d), & \text{if } 1 \le i \le d; \\ -y_{i-d} + \pi_{i-d}(x_1, \ldots, x_d), & \text{if } d+1 \le i \le 2d, \end{cases} \tag{2.8}$$

by (2.5), $z(n)$ takes the form

$$z(n+1) - z(n) = \gamma_n \left[F(z(n)) + u_n\right]. \tag{2.9}$$

The above expression is a particular case of a class of stochastic approximation algorithms studied by Benaïm 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 some appropriate conditions, the accumulation points of $\{z(n)\}_{n\ge0}$ 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, state Benaïm's theorem, and conclude the section by proving that our model satisfies the required conditions of this theorem.

2.2 Preliminaries on semiflows

Let $\Gamma \subset \mathbb{R}^m$ be a metric space and "$\mathrm{dist}(\cdot,\cdot)$" denote the metric. Let $\Phi : \mathbb{R}_+ \times \Gamma \to \Gamma$ be a continuous map. For simplicity, denote $\Phi(t, x)$ by $\Phi_t(x)$.

Definition 2.2 (Semiflow). *A semiflow on Γ is a continuous map $\Phi : \mathbb{R}_+ \times \Gamma \to \Gamma$ such that*

(i) Φ_0 is the identity on Γ, and

(ii) $\Phi_{t+s} = \Phi_t \circ \Phi_s$ for any $t, s \ge 0$.

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

$$\frac{d}{dt}\Phi_t(x) = F(\Phi_t(x)), \quad \forall\, x \in \mathbb{R}^m, \forall\, t \in \mathbb{R}_+.$$

If F is Lipschitz, then it has unique integral curves.

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

Definition 2.3 (Invariant set). *A set $A \subset \Gamma$ is called invariant if $\Phi_t(A) \subset A$ for every $t \ge 0$.*

Note that our definition of "invariant" is equivalent to the definition of "positively invariant" in some literature.

Definition 2.4 (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 Φ is the set of all equilibrium points.*

When Φ is induced by a vector field F, the equilibrium set coincides with the set of points on which F vanishes.

Definition 2.5 (Chain-recurrent point)**.** *Given $\rho, T > 0$, a point $x \in \Gamma$ is called (ρ, T)-recurrent if there are points $x_0 = x, x_1, \dots, x_{k-1}, x_k = x \in \Gamma$ and real numbers $t_0, t_1, \dots, t_{k-1} \geq T$ such that*

$$\operatorname{dist}\left(\Phi_{t_i}(x_i), x_{i+1}\right) < \rho, \quad i = 0, \dots, k-1.$$

x is said to be chain-recurrent if it is (ρ, T)-recurrent for any $\rho, T > 0$.

Let $\operatorname{CR}(\Phi)$ be the set of chain-recurrent points associated with Φ. Note that $\operatorname{CR}(\Phi)$ is closed and invariant.

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

Definition 2.6 (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 Γ is compact, $\omega(w)$ is a nonempty, compact, connected and invariant set.

Definition 2.7 (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 $t \in \mathbb{R}_+ \to L(\Phi_t(x))$ is strictly decreasing along any non-constant orbit $\Phi_t(x) \subset \Lambda$.*

2.3 A limit set theorem

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

Theorem 2.8. *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*

$$z(n+1) - z(n) = \gamma_n \left[F(z(n)) + u_n\right],$$

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 $T > 0$,*

$$\lim_{n\to\infty} \sup_k \left\{ \left\| \sum_{i=n}^{k-1} \gamma_i u_i \right\| : \sum_{i=n}^{k-1} \gamma_i \leq T \right\} = 0.$$

Then $L(\{z(n)\}_{n\geq 0})$ is a connected set chain-recurrent for the semiflow induced by F.

2.4 The random process (2.9) satisfies Theorem 2.8

First, note that $\delta 1_{x\leq\delta} + x 1_{x>\delta}$ is bounded by δ and 1, and $\varpi(x)$ is Lipschitz. Then π in (2.6) and F in (2.8) are Lipschitz. Meanwhile, $\gamma_n = 1/(n+1+d)$ satisfies

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

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

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 $\{\mathcal{F}_{n+1}\}_{n\geq 0}$

$$\mathbb{E}\left[M_{n+1}|\mathcal{F}_{n+1}\right] = \sum_{i=0}^{n} \gamma_i u_i + \mathbb{E}\left[\gamma_{n+1} u_{n+1}|\mathcal{F}_{n+1}\right] = \sum_{i=0}^{n} \gamma_i u_i = M_n.$$

Furthermore, because for any $n \geq 0$

$$\sum_{i=0}^{n} \mathbb{E}\left[\|M_{i+1} - M_i\|^2 |\mathcal{F}_{i+1}\right] \leq (2d)^2 \cdot \sum_{i=0}^{n} \gamma_{i+1}^2 \leq (2d)^2 \cdot \sum_{i\geq 0} \gamma_i^2 < \infty \text{ a.s.},$$

the sequence $\{M_n\}_{n\geq 0}$ converges to a finite random vector in $\mathbb{R}^{2d}$ almost surely (see e.g. Theorem 5.4.9 of [5]). In particular, it is a Cauchy sequence and so condition (ii) holds almost surely.

Now, in view of Theorem 2.8, we will investigate the chain-recurrent set of the semiflow generated by the following ODE

$$\begin{cases} \dfrac{du_i(t)}{dt} = -u_i(t) + \dfrac{f(v_i(t))^{-\alpha}}{\sum_{k=1}^{d} f(v_k(t))^{-\alpha}}, & i = 1, \ldots, d \\ \dfrac{dv_i(t)}{dt} = -v_i(t) + \dfrac{f(u_i(t))^{-\alpha}}{\sum_{k=1}^{d} f(u_k(t))^{-\alpha}}, & i = 1, \ldots, d \end{cases} \tag{2.10}$$

where f is a function as follows

$$f(x) = \delta 1_{x\leq\delta} + x 1_{x>\delta},\ i.e.\ f(x) = \varpi(x)^{-\frac{1}{\alpha}}. \tag{2.11}$$

We can rewrite (2.10) in vector form

$$\begin{cases} \dfrac{du(t)}{dt} = -u(t) + \pi(v(t)) \\ \dfrac{dv(t)}{dt} = -v(t) + \pi(u(t)) \end{cases} \quad \text{or} \quad \frac{d\Xi(t)}{dt} = F(\Xi(t)) \tag{2.12}$$

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

Before moving to the proof of Theorem 1.1, we will prove a simple fact regarding (2.10).

Proposition 2.9. *The domain D is invariant under Φ, the semiflow induced by (2.10).*

Proof. Suppose $(u, v) \in \partial D$. Without loss of generality, we can assume that there exists some $i \in V$ such that $u_i = 0$. Then by (2.10), 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. This completes the proof. □

3 Proof of the main result

By Theorem 2.8, the limit set of $\{z(n)\}_{n\geq 0}$ is contained in the chain recurrent set, and so the first step to prove Theorem 1.1 is to characterize chain-recurrent set for our specific semiflow induced by (2.10). Recall U defined in Notation 2.1. We will conclude the proof of Theorem 1.1 by showing that $\{z(n)\}_{n\geq 0}$ has probability zero to converge to the isolated unstable equilibrium (U, U).

3.1 Chain recurrent set

3.1.1 Lyapunov function

We characterize the chain-recurrent set $\mathrm{CR}(\Phi)$ by introducing a Lyapunov function

$$L(u,v) = \sum_{i=1}^{d} u_i v_i, \quad (u,v) \in D. \tag{3.1}$$

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 Φ where $x = (u_1(0), \ldots, u_d(0), v_1(0), \ldots, v_d(0))$. Then

$$\begin{aligned}
&\frac{d}{dt}(L(\Phi_t(x))) \\
= \;& \sum_{i=1}^{d} v_i(t) \frac{du_i(t)}{dt} + \sum_{i=1}^{d} u_i(t) \frac{dv_i(t)}{dt} \\
= \;& \sum_{i=1}^{d} v_i \left(-u_i + \frac{f(v_i)^{-\alpha}}{\sum_{k=1}^{d} f(v_k)^{-\alpha}} \right) + \sum_{i=1}^{d} u_i \left(-v_i + \frac{f(u_i)^{-\alpha}}{\sum_{k=1}^{d} f(u_k)^{-\alpha}} \right) \\
= \;& -2 \sum_{i=1}^{d} u_i v_i + \frac{\sum_{i=1}^{d} u_i f(u_i)^{-\alpha}}{\sum_{k=1}^{d} f(u_k)^{-\alpha}} + \frac{\sum_{i=1}^{d} v_i f(v_i)^{-\alpha}}{\sum_{k=1}^{d} f(v_k)^{-\alpha}}.
\end{aligned} \tag{3.2}$$

Notice that the right hand side of (3.2) depends on t only through dependence on $u_i(t)$ and $v_i(t)$. We have the following lemma about (3.2), which confirms that $L(u,v)$ is a Lyapunov function for a large subset of the domain D according to Definition 2.7.

Lemma 3.1. *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 δ, s.t. when $\alpha \geq \alpha(d)$*

$$\left. \frac{d}{dt}(L(\Phi_t(x))) \right|_{(u,v)} \leq 0, \quad \forall (u,v) \in D^\delta, \tag{3.3}$$

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

To prove Lemma 3.1, we need several other lemmas. Recall that $V = \{1, \ldots, d\}$.

Lemma 3.2. *When $\alpha > d-2$, U is a local minimum of the function $g : \mathring{\Delta} \to \mathbb{R}$ defined as*

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

In particular, $g(U) = 0$.

Proof. Define a function G on $\mathbb{R}_+^d$:

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

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

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

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

$$G(w_1, \ldots, w_d) = \frac{2 w_d}{\sum_{i=1}^{d} w_i} - \frac{d w_d^2}{(\sum_{i=1}^{d} w_i)^2} - \frac{1}{\sum_{i=1}^{d} w_i} \cdot \frac{\sum_{i=1}^{d} w_i^{-\alpha}}{\sum_{i=1}^{d} w_i^{-(\alpha+1)}}. \tag{3.4}$$

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

First, by direct calculation, we can check that $G(w_1, \ldots, w_d)$ has zero gradient at W, i.e. $\nabla G|_W = 0$. Hence, W is a critical point of $G(w_1, \ldots, w_d)$.

Further, we will prove that $G(w_1, \ldots, w_d)$ is convex along all the other directions except the diagonal. We calculate H, the Hessian matrix of $G(w_1, \ldots, w_d)$ at W:

$$H = \frac{2}{d^2w^2}\begin{pmatrix} -\frac{\alpha+2}{d}+\alpha+1 & \ldots & -\frac{\alpha+2}{d} & 1-\frac{\alpha+2}{d} \\ \vdots & \ddots & \vdots & \vdots \\ -\frac{\alpha+2}{d} & \ldots & -\frac{\alpha+2}{d}+\alpha+1 & 1-\frac{\alpha+2}{d} \\ 1-\frac{\alpha+2}{d} & \ldots & 1-\frac{\alpha+2}{d} & -\frac{\alpha+2}{d}+\alpha+1+2-d \end{pmatrix}.$$

Let

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

and hence $H = (2/(d^2w^2))P$. We want to calculate the eigenvalues of P first. Let

$$Q = \begin{pmatrix} -\frac{\alpha+2}{d} & -\frac{\alpha+2}{d} & \ldots & -\frac{\alpha+2}{d} & 1-\frac{\alpha+2}{d} \\ -\frac{\alpha+2}{d} & -\frac{\alpha+2}{d} & -\frac{\alpha+2}{d} & \ldots & 1-\frac{\alpha+2}{d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\frac{\alpha+2}{d} & \ldots & -\frac{\alpha+2}{d} & -\frac{\alpha+2}{d} & 1-\frac{\alpha+2}{d} \\ 1-\frac{\alpha+2}{d} & \ldots & \ldots & 1-\frac{\alpha+2}{d} & -\frac{\alpha+2}{d}+2-d \end{pmatrix}$$

and hence $P = (\alpha+1)I + Q$, where I is the identity matrix. By direct calculation, we can have the eigenvalues of Q

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

Then shifting Q's eigenvalues by $\alpha+1$, we get the eigenvalues of P

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

Finally, we derive the eigenvalues of H

$$\lambda_1^H = \ldots = \lambda_{d-2}^H = \frac{2(\alpha+1)}{d^2w^2}, \lambda_{d-1}^H = 0, \lambda_d^H = \frac{2(\alpha+2-d)}{d^2w^2}.$$

Thus when $\alpha > d-2$, one of H's eigenvalues is zero and all the others are strictly positive. It is easy to check that the sum of each row of H is zero, which means

$$H\left(1, \ldots, 1\right)^T = 0 \cdot \left(1, \ldots, 1\right)^T.$$

That is, the diagonal is an eigenvector associated with H's zero eigenvalue. This proves 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. □

Keeping the notations of Lemma 3.2, we have the following lemma.

Lemma 3.3. *For any fixed positive integer $d \geq 3$, there exists some $\alpha_0(d)$, such that when $\alpha > \alpha_0(d)$, U is the global minimum of $g(u_1, \ldots, u_d)$.*

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

$$2 \min_{i \in V} u_i - d \left(\min_{i \in V} u_i \right)^2 \geq \frac{\sum_{i=1}^d u_i^{-\alpha}}{\sum_{i=1}^d u_i^{-(\alpha+1)}}, \tag{3.5}$$

with equality if and only if $u = U$.

We will divide the proof of (3.5) into two cases:

(1) u is in a neighborhood of U;

(2) u is bounded away from U. Equivalently, for some fixed $0 < \kappa < 1$, u satisfies $\min_{i \in V} u_i < \kappa/d$.

Case (1) directly follows from Lemma 3.2.

To prove case (2), first we use the minimum coordinates of u to bound the right hand side of (3.5) from above. More precisely, for fixed d and α, we will prove that for any $u \in \overset{\circ}{\Delta}$, the following holds

$$\frac{\sum_{i=1}^d u_i^{-\alpha}}{\sum_{i=1}^d u_i^{-(\alpha+1)}} \leq d^{1/(\alpha+1)} \min_{i \in V} u_i. \tag{3.6}$$

Without loss of generality, we can assume $u_d = \min_{i \in V} u_i$. Then if letting $a_i = \min_{i \in V} u_i / u_i = u_d / u_i \in (0, 1]$, (3.6) is equivalent to the following inequality with $a_i \in (0,1]\,(i = 1, \dots, d-1)$

$$\frac{1 + \sum_{i=1}^{d-1} a_i^{\alpha}}{1 + \sum_{i=1}^{d-1} a_i^{\alpha+1}} \leq d^{1/(\alpha+1)}. \tag{3.7}$$

To prove (3.7), 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

$$\begin{aligned} \frac{1 + \sum_{i=1}^{d-1} a_i^{\alpha}}{1 + \sum_{i=1}^{d-1} a_i^{\alpha+1}} &\leq \frac{d^{1/(\alpha+1)} \left(1 + \sum_{i=1}^{d-1} a_i^{\alpha+1} \right)^{\alpha/(\alpha+1)}}{1 + \sum_{i=1}^{d-1} a_i^{\alpha+1}} \\ &= \frac{d^{1/(\alpha+1)}}{\left(1 + \sum_{i=1}^{d-1} a_i^{\alpha+1} \right)^{1/(\alpha+1)}} < d^{1/(\alpha+1)}, \end{aligned}$$

thus proving (3.7) and also (3.6). Notice that when $\alpha \geq \log d / \log(2 - \kappa) - 1$, for any $u_d \in (0, \kappa/d)$ the following inequality holds

$$d^{1/(\alpha+1)} \cdot u_d < 2u_d - du_d^2, \quad i.e. \quad d^{1/(\alpha+1)} < 2 - du_d. \tag{3.8}$$

Then (3.6) and (3.8) together imply that when $\alpha \geq \log d / \log(2 - \kappa) - 1$, for any u satisfying $\min_{i \in V} u_i < \kappa/d$

$$2 \min_{i \in V} u_i - d \left(\min_{i \in V} u_i \right)^2 > \frac{\sum_{i=1}^d u_i^{-\alpha}}{\sum_{i=1}^d u_i^{-(\alpha+1)}}. \tag{3.9}$$

This finishes the proof of case (2).

Finally, we need to combine the two cases together. From case (1), for fixed d and $\alpha > d-2$, there exists a neighborhood of the uniform distribution $\mathcal{N}(U,\epsilon_\alpha)$ s.t. for any $u \in \mathcal{N}(U,\epsilon_\alpha)$, (3.5) holds. For any $u \neq U$ in $\mathring{\Delta}$, since $\sum_{i=1}^d u_i^{-\alpha} / \sum_{i=1}^d u_i^{-(\alpha+1)}$ is a decreasing function[2] in α, $g(u_1,\ldots,u_d)$ in Lemma 3.2 is an increasing function in α. This allows us to take some common neighborhood $\mathcal{N}(U,\epsilon_d) = \bigcap_{\alpha > d-1} \mathcal{N}(U,\epsilon_\alpha)$ just depending on d such that (3.5) holds. Take some $\kappa = \kappa(d) < 1$ such that

$$\left\{ u \in \mathring{\Delta} : \min_{i\in V} u_i < \frac{\kappa}{d} \right\} \cup \mathcal{N}(U,\epsilon_d) = \mathring{\Delta}.$$

Then let

$$\alpha_0(d) = \max\left\{ d-1,\ \frac{\log d}{\log(2-\kappa)} - 1 \right\}.$$

When $\alpha > \alpha_0(d)$, the above two cases combined imply that (3.5) holds for any $u \in \mathring{\Delta}$. $\square$

Lemma 3.4. *For any fixed positive integer $d \geq 3$, there exists some $\alpha_0(d)$, such that when $\alpha > \alpha_0(d)$, for any $(u,v) = (u_1,\ldots,u_d,v_1,\ldots,v_d) \in \mathring{D}$ the following holds*

$$2\sum_{i=1}^d u_i v_i \geq \frac{\sum_{i=1}^d u_i^{-\alpha}}{\sum_{i=1}^d u_i^{-(\alpha+1)}} + \frac{\sum_{i=1}^d v_i^{-\alpha}}{\sum_{i=1}^d v_i^{-(\alpha+1)}}, \tag{3.10}$$

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

Proof. First for any fixed $u, v \in \mathring{\Delta}$, we will bound the left hand side of (3.10) from below by a function of the minimum coordinates of u and v. More precisely, we construct two $d-$dimensional vectors u' and v' by the minimum coordinates of u and v

$$u' = \left(\min_{i\in V} u_i, \ldots, \min_{i\in V} u_i, 1-(d-1)\min_{i\in V} u_i \right)$$

and

$$v' = \left(1-(d-1)\min_{i\in V} v_i, \min_{i\in V} v_i, \ldots, \min_{i\in V} v_i \right),$$

and then we will prove that

$$\sum_{i=1}^d u_i v_i \geq \sum_{i=1}^d u_i' v_i' = \min_{i\in V} u_i + \min_{i\in V} v_i - d\cdot \min_{i\in V} u_i \min_{i\in V} v_i. \tag{3.11}$$

By the Rearrangement inequality, it suffices to prove (3.11) for any $(u,v) \in \mathring{D}$ satisfying $u_1 \leq \ldots \leq u_d$ and $v_1 \geq \ldots \geq v_d$. For such u and v, we have

$$\begin{aligned}
\sum_{i=1}^d u_i v_i &= u_1 v_1 + \sum_{i=2}^d u_i v_d + \sum_{i=2}^d u_i (v_i - v_d) \\
&\geq u_1 v_1 + \sum_{i=2}^d u_i v_d + \sum_{i=2}^d u_1 (v_i - v_d) \\
&= u_1(1-(d-1)v_d) + \sum_{i=2}^d u_i v_d \\
&= \sum_{i=1}^d u_i v_i' \geq \sum_{i=1}^d u_i' v_i',
\end{aligned}$$

[2]This can be proved by looking at the derivative.

where the last step is obtained by repeating the argument in the previous steps. Thus we have proved (3.11).

By Lemma 3.3, when $\alpha > \alpha_0(d)$, for any $u, v \in \mathring{\Delta}$

$$2\min_{i\in V} u_i - d\left(\min_{i\in V} u_i\right)^2 + 2\min_{i\in V} v_i - d\left(\min_{i\in V} v_i\right)^2 \geq \frac{\sum_{i=1}^d u_i^{-\alpha}}{\sum_{i=1}^d u_i^{-(\alpha+1)}} + \frac{\sum_{i=1}^d v_i^{-\alpha}}{\sum_{i=1}^d v_i^{-(\alpha+1)}} \tag{3.12}$$

Observe that the following elementary inequality holds

$$2\left(\min_{i\in V} u_i + \min_{i\in V} v_i - d\min_{i\in V} u_i \min_{i\in V} v_i\right) \geq 2\min_{i\in V} u_i - d\left(\min_{i\in V} u_i\right)^2 + 2\min_{i\in V} v_i - d\left(\min_{i\in V} v_i\right)^2 \tag{3.13}$$

Then it follows from (3.12) and (3.13) that

$$2\left(\min_{i\in V} u_i + \min_{i\in V} v_i - d\min_{i\in V} u_i \min_{i\in V} v_i\right) \geq \frac{\sum_{i=1}^d u_i^{-\alpha}}{\sum_{i=1}^d u_i^{-(\alpha+1)}} + \frac{\sum_{i=1}^d v_i^{-\alpha}}{\sum_{i=1}^d v_i^{-(\alpha+1)}}. \tag{3.14}$$

Finally, (3.11) and (3.14) together imply (3.10).

It is also easy to check that the equality in (3.10) holds if and only if $u = v = U$. □

Proof of Lemma 3.1. Recall that $f(x) = \delta 1_{x\leq\delta} + x1_{x>\delta}$. 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 $\min_{i\in V} f(v_i) = \delta$ and $\min_{i\in V} f(u_i) > 2\delta$;

(3) $\min_{i\in V} f(u_i) = \delta$ and $\min_{i\in V} f(v_i) \leq 2\delta$, or the symmetric case $\min_{i\in V} f(v_i) = \delta$ and $\min_{i\in V} f(u_i) \leq 2\delta$.

Let's 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 (3.3), by (3.2), it is equivalent to prove (3.10). Then by Lemma 3.4, case (1) follows if $\alpha(d) > \alpha_0(d) + 1$. Notice that in case (1), by Lemma 3.4, when $\alpha(d) > \alpha_0(d) + 1$, $(u, v) = (U, U)$ is the only point where $\frac{d}{dt}(L(\Phi_t(x)))\big|_{(u,v)} \leq 0$ can hold with equality.

Now we prove case (2). We only prove the case $\min_{i\in V} f(u_i) = \delta$ and $\min_{i\in V} f(v_i) > 2\delta$. By (3.2),

$$\begin{aligned}
\frac{d}{dt}(L(\Phi_t(x)))\bigg|_{(u,v)} &= -2\sum_{i=1}^d u_i v_i + \frac{\sum_{i=1}^d u_i f(u_i)^{-\alpha}}{\sum_{k=1}^d f(u_k)^{-\alpha}} + \frac{\sum_{i=1}^d v_i f(v_i)^{-\alpha}}{\sum_{k=1}^d f(v_k)^{-\alpha}} \\
&\leq -2\sum_{i=1}^d u_i v_i + \frac{\sum_{i=1}^d f(u_i) f(u_i)^{-\alpha}}{\sum_{k=1}^d f(u_k)^{-\alpha}} + \frac{\sum_{i=1}^d f(v_i) f(v_i)^{-\alpha}}{\sum_{k=1}^d f(v_k)^{-\alpha}} \\
&\leq -2\sum_{i=1}^d u_i v_i + d^{1/\alpha}\left(\min_{i\in V} f(u_i) + \min_{i\in V} f(v_i)\right),
\end{aligned} \tag{3.15}$$

where the last step is by (3.6), 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 assumptions $\min_{i\in V} f(u_i) = \delta$ and $\min_{i\in V} f(v_i) > 2\delta$ 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,

$$\begin{aligned}\frac{d}{dt}(L(\Phi_t(x)))\Big|_{(u,v)} &\leq -2\sum_{i=1}^{d} u_i v_i + d^{1/\alpha}\left(\min_{i\in V} f(u_i) + \min_{i\in V} f(v_i)\right)\\ &\leq -2\min_{i\in V} f(v_i) + d^{1/\alpha}\left(\delta + \min_{i\in V} f(v_i)\right)\\ &= -\left(2 - d^{1/\alpha}\right)\min_{i\in V} f(v_i) + d^{1/\alpha}\delta\\ &\leq -\left(2 - d^{1/\alpha}\right)2\delta + d^{1/\alpha}\delta = -\left(4 - 3d^{1/\alpha}\right)\delta.\end{aligned}$$

Thus if choosing $\alpha > \log d/\log\frac{4}{3}$ such that $4 - 3d^{1/\alpha} > 0$, we obtain $\frac{d}{dt}(L(\Phi_t(x)))\big|_{(u,v)} < 0$ in case (2).

In case (3), we just prove the case $\min_{i\in V} f(u_i) = \delta$ and $\min_{i\in V} f(v_i) \leq 2\delta$. First we can choose $\alpha > \log_2 d$ such that $\frac{3}{2}d^{1/\alpha} < 3$. Then by the definition of D^δ,

$$L(u,v) \geq 3\delta > \frac{3}{2}d^{1/\alpha}\delta, \quad \forall (u,v) \in D^\delta. \tag{3.16}$$

(3.15) and (3.16) together imply

$$\frac{d}{dt}(L(\Phi_t(x)))\Big|_{(u,v)} < -2\sum_{i=1}^{d} u_i v_i + 3d^{1/\alpha}\delta < 0, \tag{3.17}$$

establishing case (3).

To sum up, taking

$$\alpha(d) = \max\left\{\alpha_0(d) + 1,\ \log d/\log\frac{4}{3},\ \log_2 d\right\} + 1,$$

we prove the lemma. □

3.1.2 The main lemma

Now it comes to the main lemma to characterize the chain recurrent set $\mathrm{CR}(\Phi)$ for our specific semiflow Φ.

Lemma 3.5. *Let $S^\delta = \{(u,v) \in D : \sum_{i=1}^{d} u_i v_i \leq 4\delta\}$. Then $\mathrm{CR}(\Phi) \subset S^\delta \cup (U,U)$.*

Proof. Let $\zeta_0 = 3.5\delta$, $\zeta_1 = 4\delta$ and $\zeta_2 = 1$. Define

$$M_j = \{(x,y) \in D : L(x,y) \leq \zeta_j\}, \quad j = 0,1,2.$$

Note that $M_1 = S^\delta, M_2 = D$, and $(U,U) \in M_2 \setminus M_1$ when δ is small enough. By Lemma 3.1 and Proposition 2.9, M_j $(j = 0,1,2)$ are compact invariant sets. Clearly, the lemma will follow once we prove:

(a) $CR_1 := \mathrm{CR}(\Phi) \cap M_1$ and $CR_2 := \mathrm{CR}(\Phi) \cap (M_2 \setminus M_1)$ are invariants sets;

(b) $CR_2 = (U,U)$.

Let's prove (a).

By the invariance of $\mathrm{CR}(\Phi)$ and M_1, it is clear that CR_1 is invariant.

Now we prove by contradiction that CR_2 is invariant. Suppose CR_2 is not invariant. Then there exists some $z \in CR_2$, s.t. $\Phi_{T_0}(z) \in M_1$ for some $T_0 > 0$. Then by Lemma 3.1 and compactness of $\overline{M_1 \backslash M_0}$, there exists some $T_1 > T_0$, such that for all $t > T_1$

$$L(\Phi_t(z)) < \zeta_0, \; i.e. \; \Phi_t(z) \in M_0. \tag{3.18}$$

Also by Lemma 3.1 and compactness of $\overline{M_1 \backslash M_0}$, there exists some $T_2 > 0$, such that for all $t > T_2$

$$\Phi_t(M_1) \subset M_0. \tag{3.19}$$

Let $\rho_0 = \mathrm{dist}\left(M_0, \overline{D \backslash M_1}\right) > 0$ and $T = \max\{T_1, T_2\}$. By the assumption $z \in CR_2 \subset CR(\Phi)$, 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

$$\mathrm{dist}\left(\Phi_{t_i}(z_i), z_{i+1}\right) < \rho_0, \quad i = 0, \ldots, k-1. \tag{3.20}$$

By (3.18), $\Phi_{t_0}(z_0) = \Phi_{t_0}(z) \in M_0$ and so, by (3.20), $z_1 \in M_1$. By induction, we claim that $z_1, z_2, \ldots, z_k \in M_1$. Indeed, if $z_i \in M_1$, by (3.19), $\Phi_{t_i}(z_i) \in M_0$, and then by (3.20), $z_{i+1} \in M_1$. In particular, $z_k = z \in M_1$, which contradicts the assumption $z \in M_2 \setminus M_1$. Hence, $\Phi_t(CR_2) \subset M_2 \setminus M_1$ for all $t \geq 0$. By invariance of $\mathrm{CR}(\Phi)$, 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. Let

$$L^\infty = \lim_{t \to +\infty} L(\Phi_t(z)).$$

Then for any $p \in \omega(z)$ (the omega limit set of z), $L(p) = L^\infty$. 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, $\omega(z) = (U, U)$. Further, we will prove by contradiction that $z = (U, U)$. 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 $\epsilon > 0$, s.t. $L(z) > 1/d + \epsilon$. Let

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

By Lemma 3.1 and compactness of F, there exists some T', such that for all $t > T'$, $\Phi_t(F) \subset E$. Let $\rho_1 = \mathrm{dist}\,(E, F) > 0$. By the assumption that $z \in CR_2 \subset CR(\Phi)$, z is (ρ_1, T')-recurrent. Then by a similar argument as the proof of (a), letting E play the role of M_0, we can get the desired contradiction. □

3.2 Non-convergence to unstable equilibrium

By Theorem 2.8 and Lemma 3.5,

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

Since S^δ and (U, U) are disconnected, we can finish the proof of Theorem 1.1 by proving the following lemma.

Lemma 3.6. *When $\alpha > 1$, $z(n)$ in (1.4) satisfies*

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

Lemma 3.6 is an easy application of a theorem due to Pemantle [12] which, to our purposes, is stated as

Theorem 3.7. *[12, Theorem 1] Let $z(n)$ be a stochastic process satisfying*

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

with $\mathbb{E}(u_n|\mathcal{F}_n) = 0$. Assume that $z(n)$ always remains in a bounded domain D. Let p be any point in $\mathring{D}$ with $F(p) = 0$, and $\mathcal{N}$ be a neighborhood of p. 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 an unstable critical point,

(2) $\mathbb{E}((u_n \cdot \theta)^+|\mathcal{F}_n) \geq c_1$ for every unit vector $\theta \in TD$ (see the definition of TD in Notation 2.1),

(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 that 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 $z(n)$ in (1.4) satisfies the conditions of Theorem 3.7 with $p = (U,U)$. First it is easy to check that p is a critical point of the vector field F in (2.8), i.e. $F(U,U) = 0$. Before proving that p is unstable, we need to introduce a formal definition.

Definition 3.8 (attracting/unstable point)**.** *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 point.

Lemma 3.9. *1. When $\alpha > 1$, (U,U) is an unstable point of F in (2.8).*

2. When $\alpha < 1$, (U,U) is an attracting point of F in (2.8).

Proof. In a neighborhood of (U,U), F has the Taylor expansion

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

where $DF|_p$ is the Jacobian matrix at p and w is some vector in a neighborhood of 0 (a 2d dimensional vector). By direct calculation, $DF|_p$ has expression

$$DF|_p = \left(\begin{array}{cccc|cccc} -1 & 0 & \dots & 0 & -\alpha+\frac{\alpha}{d} & \frac{\alpha}{d} & \dots & \frac{\alpha}{d} \\ 0 & -1 & 0 & \dots & \frac{\alpha}{d} & -\alpha+\frac{\alpha}{d} & \frac{\alpha}{d} & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & \frac{\alpha}{d} & \dots & \frac{\alpha}{d} & -\alpha+\frac{\alpha}{d} \\ \hline -\alpha+\frac{\alpha}{d} & \frac{\alpha}{d} & \dots & \frac{\alpha}{d} & -1 & 0 & \dots & 0 \\ \frac{\alpha}{d} & -\alpha+\frac{\alpha}{d} & \frac{\alpha}{d} & \dots & 0 & -1 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha}{d} & \dots & \frac{\alpha}{d} & -\alpha+\frac{\alpha}{d} & 0 & \dots & 0 & -1 \end{array}\right).$$

To get the eigenvalues of $DF|_p$, we need to solve an equation of matrix's determinant:

$$\left| DF|_p - \lambda I_{2d\times 2d} \right| = 0. \tag{3.21}$$

Notice that $DF|_p$ has the same upper-right and lower-left block matrix as marked in the expression. Let B denote this $d \times d$ block matrix. Because the sum of B's each row is zero, B has a zero eigenvalue and hence zero determinant. Then one can easily check that $\lambda = -1$ is a solution to (3.21). Now assume $\lambda \neq -1$. Then we can apply a formula of Schur complement, and derive from (3.21) that

$$\left| B^2 - (\lambda+1)^2 I_{d\times d} \right| = 0, \tag{3.22}$$

where $I_{d\times d}$ is the d dimensional identity matrix. (3.22) is equivalent to

$$|B - (\lambda+1) I_{d\times d}| \cdot |B + (\lambda+1) I_{d\times d}| = 0. \tag{3.23}$$

Under the assumption $\lambda \neq -1$, we can easily solve (3.23): $\lambda = -1 \pm \alpha$. Hence the eigenvalues of $DF|_p$ (without counting multiplicities) are $-1, -1 \pm \alpha$. So when $\alpha > 1$, $DF|_p$ has a positive eigenvalue $-1+\alpha$; When $\alpha < 1$, all of its eigenvalues are strictly negative. This completes the proof of the lemma. ☐

Clearly, u_n in (2.7) satisfies condition (3) of Theorem 3.7. It remains to check condition (2), which is the statement of the following lemma.

Lemma 3.10. *In a small neighborhood of $p = (U, U)$, there exists some constant $c_1 > 0$, s.t. $\mathbb{E}((u_n \cdot \theta)^+ | \mathcal{F}_n) \geq c_1$ for every unit vector $\theta = (\theta_k)_{1\leq k \leq 2d} \in TD$.*

Proof. For any fixed $i, j \in V$, conditioning on the event that $X_{n+1} = i, Y_{n+1} = j$,

$$\begin{aligned} u_n \cdot \theta &= (1 - \pi_i(y(n)))\theta_i - \sum_{m\neq i, m\in V} \pi_m(y(n))\theta_m \\ &\quad + (1 - \pi_j(x(n)))\theta_{j+d} - \sum_{s\neq j, s\in V} \pi_s(x(n))\theta_{s+d} \\ &= \theta_i + \theta_{j+d} - \sum_{m\in V} \pi_m(y(n))\theta_m - \sum_{s\in V} \pi_s(x(n))\theta_{s+d}. \end{aligned} \tag{3.24}$$

Now we will prove that for any unit vector $\theta \in TD$, its maximum coordinate is bounded from below by a positive number, and more precisely,

$$\max_{1\leq k\leq 2d} \theta_k \geq \frac{1}{2d(d-1)}. \tag{3.25}$$

Observe that θ as a unit vector ($\|\theta\| = 1$) always satisfies

$$\max_{1\leq k\leq 2d} |\theta_k| \geq \frac{1}{2d}, \tag{3.26}$$

but it does not necessarily satisfy

$$\max_{1\leq k\leq 2d} \theta_k \geq \frac{1}{2d}. \tag{3.27}$$

If θ satisfies (3.27), (3.25) naturally holds for this θ since $1/(2d) > 1/[2d(d-1)]$. If θ doesn't satisfy (3.27), by (3.26), it must hold that $\min_{1\leq k\leq 2d} \theta_k \leq -1/(2d)$. 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 again implies (3.25).

By (3.25), without loss of generality, we can assume

$$\theta_1 = \max_{1\le k\le 2d} \theta_k \ge \frac{1}{2d(d-1)}. \tag{3.28}$$

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

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 both $\sum_{m\in V} \pi_m(y(n))\theta_m$ and $\sum_{s\in V} \pi_s(x(n))\theta_{s+d}$ in (3.24) are close to zero. Therefore, by (3.24),

$$\begin{aligned}\mathbb{E}((u_n\cdot\theta)^+|\mathcal{F}_n) &\ge \mathbb{P}(X_{n+1}=1, Y_{n+1}=j_0|\mathcal{F}_n)\theta_1\\ &= \mathbb{P}(X_{n+1}=1|\mathcal{F}_n)\mathbb{P}(Y_{n+1}=j_0|\mathcal{F}_n)\theta_1.\end{aligned}$$

Again by the 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 (3.28), 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. □

Finally, we can apply Theorem 3.7, obtaining Lemma 3.6. This also completes the proof of Theorem 1.1.

4 Further problems

This paper is a starting point to understand the behavior of multi-particle repelling random walks. The general question remains widely open. The dynamical approach should still work, but the corresponding dynamical system only gets more complex and harder to analyze.

Regarding the model we just studied, we conjecture that

Conjecture 4.1. *For any positive integer $d\ge 3$, $\alpha_c = 1$ is critical. Namely,*

1. *when $\alpha > 1$, there exists some constant $c = c(\alpha, d)$ depending on α and d, such that the following holds*

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

2. *when $0<\alpha<1$, the following holds*

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

Another problem of interest is

Problem 4.2. When $\delta = 0$ in (1.2) and (1.3), is it possible to derive a similar result as Theorem 1.1?

Notice that when $\delta = 0$, the vector field F is not well defined on the boundary of D. Particularly, F won't be continuous at the boundary, and hence Theorem 2.8 and then the proof of Theorem 1.1 are invalid in this case.

References

[1] Angel, O., Crawford, N. and Kozma, G.: Localization for linearly edge reinforced random walks. *Duke Math. J.* **163**, (2014), 889–921. MR-3189433

[2] Benaïm, M.: A dynamical system approach to stochastic approximations. *SIAM J. Control Optim.* **34**, (1996), 437–472. MR-1377706

[3] Benaïm, M., Raimond, O. and Schapira, B.: Strongly reinforced vertex-reinforced-random-walk on the complete graph. *ALEA Lat. Am. J. Probab. Math. Stat.* **10**, (2013), 767–782. MR-3125746

[4] Benveniste, A., Métivier, M. and Priouret, P.: Adaptive algorithms and stochastic approximations. Translated from the French by Stephen S. Wilson. *Springer-Verlag*, Berlin, 1990. xii+365 pp. MR-1082341

[5] Durrett, R.: Probability: theory and examples. Fourth edition. Cambridge Series in Statistical and Probabilistic Mathematics. *Cambridge University Press*, Cambridge, 2010. x+428 pp. MR-2722836

[6] Kiefer, J. and Wolfowitz, J.: Stochastic estimation of the maximum of a regression function. *Ann. Math. Statistics* **23**, (1952), 462–466. MR-0050243

[7] Kovchegov, Y.: Multi-particle processes with reinforcements. *J. Theoret. Probab.* **21**, (2008), 437–448. MR-2391254

[8] Kushner, H. J. and Clark, D. S.: Stochastic approximation methods for constrained and unconstrained systems. Applied Mathematical Sciences, 26. *Springer-Verlag*, New York-Berlin, 1978. x+261 pp. MR-0499560

[9] Limic, V.: Attracting edge property for a class of reinforced random walks. *Ann. Probab.* **31**, (2003), 1615–1654. MR-1989445

[10] Limic, V. and Tarrès, P.: Attracting edge and strongly edge reinforced walks. *Ann. Probab.* **35**, (2007), 1783–1806. MR-2349575

[11] Ljung, L.: Analysis of recursive stochastic algorithms. *IEEE Trans. Automatic Control* **AC-22**, (1977), 551–575. MR-0465458

[12] Pemantle, R.: Nonconvergence to unstable points in urn models and stochastic approximations. *Ann. Probab.* **18**, (1990), 698–712. MR-1055428

[13] Pemantle, R.: A survey of random processes with reinforcement. *Probab. Surv.* **4**, (2007), 1–79. MR-2282181

[14] Pemantle, R. and Volkov, S.: Vertex-reinforced random walk on $\mathbb{Z}$ has finite range. *Ann. Probab.* **27**, (1999), 1368–1388. MR-1733153

[15] Robbins, H. and Monro, S.: A stochastic approximation method. *Ann. Math. Statistics* **22**, (1951), 400–407. MR-0042668

[16] Tarrès, P.: Vertex-reinforced random walk on $\mathbb{Z}$ eventually gets stuck on five points. *Ann. Probab.* **32**, (2004), 2650–2701. MR-2078554

Acknowledgments. During the preparation of this manuscript, the author was a student at the Weizmann Institute of Science, and this work was supported by the ISF. The author thanks Itai Benjamini for his very stimulating question 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 the discussions on dynamical systems.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])
- Non profit, sponsored by IMS[2], BS[3], PKP[4]
- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]OJS: Open Journal Systems `http://pkp.sfu.ca/ojs/`
[2]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[3]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[4]PK: Public Knowledge Project `http://pkp.sfu.ca/`
[5]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[6]IMS Open Access Fund: `http://www.imstat.org/publications/open.htm`