Self-Repellent Random Walks on General Graphs - Achieving Minimal Sampling Variance via Nonlinear Markov Chains

Vishwaraj Doshi [1] **Jie Hu** [2] **Do Young Eun** [2]

Abstract

We consider random walks on discrete state spaces, such as general undirected graphs, where the random walkers are designed to approximate a target quantity over the network topology via sampling and neighborhood exploration in the form of Markov chain Monte Carlo (MCMC) procedures. Given any Markov chain corresponding to a target probability distribution, we design a *self-repellent random walk* (SRRW) which is less likely to transition to nodes that were highly visited in the past, and more likely to transition to seldom visited nodes. For a class of SRRWs parameterized by a positive real α, we prove that the empirical distribution of the process converges almost surely to the the target (stationary) distribution of the underlying Markov chain kernel. We then provide a central limit theorem and derive the exact form of the arising asymptotic co-variance matrix, which allows us to show that the SRRW with a stronger repellence (larger α) always achieves a smaller asymptotic covariance, in the sense of Loewner ordering of co-variance matrices. Especially for SRRW-driven MCMC algorithms, we show that the decrease in the asymptotic sampling variance is of the order $O(1/\alpha)$, eventually going down to zero. Finally, we provide numerical simulations complimentary to our theoretical results, also empirically testing a version of SRRW with α increasing in time to combine the benefits of smaller asymptotic variance due to large α, with empirically observed faster mixing properties of SRRW with smaller α.

[1]IQVIA Inc., Plymouth Meeting, USA. [2]Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, USA.. Correspondence to: Do Young Eun <dyeun@ncsu.edu>.

Proceedings of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023. Copyright 2023 by the author(s).

1. Introduction

Random walk based techniques are a staple in statistics and learning theory. Markov chains such as the Metropolis Hastings random walk, designed to achieve any given target probability distribution as its stationary measure, are widely used as Markov chain Monte Carlo (MCMC) samplers and in distributed optimization via stochastic gradient descent (Sun et al., 2018; Hu et al., 2022). The local nature of the information required to compute state transition probabilities means that the algorithms scale well and are robustly implementable over state spaces such as large graphs/networks with general topologies. However, classic Markov chains can often be victims of limitations set by the underlying topology of the state space (communication matrix or adjacency matrix of the underlying network structure) leading to correlated samples which can negatively affect the estimator performance. It has also been well established that the time-reversibility requirement for the classical MCMC samplers is one of the causes for their slow convergence (see Andrieu & Livingstone, 2021, Section 1). One way in which this problem has been approached in the literature is via construction of non-reversible versions of the base Markov chain (Diaconis et al., 2000; Turitsyn et al., 2011; Chen & Hwang, 2013; Ma et al., 2016; Thin et al., 2020), which is often done by inducing some form of *non-backtracking* behaviour, that is, avoiding states most recently visited by the random walker (Alon et al., 2007). This involves the random walker interacting with some of its own past history, and has been shown to possess better efficiency than the original *base* Markov chain in the sense of the MCMC estimator achieving a smaller asymptotic variance (Neal, 2004; Lee et al., 2012). Since these non-backtracking based methods only utilize the most *recent* history of the random walker and are still provably more efficient, it is natural to consider the design of protocols where the random walker interacts with its *entire* past history to speed up its diffusion and increase its sampling efficiency, especially for sampling over discrete state spaces. This is the approach taken in our paper.

Let $\mathcal{G}(\mathcal{N}, \mathcal{E})$ be an undirected, connected graph where $\mathcal{N} \triangleq \{1, \cdots, N\}$ denotes the set of nodes and $\mathcal{E}$ denotes the set of edges, where we say $(i, j) \in \mathcal{E}$ if there is an edge between

nodes $i, j \in \mathcal{N}$. We use $\mathbf{A} = [a_{ij}]_{i,j\in\mathcal{N}}$ to represent the adjacency matrix of the graph, where $a_{ij} > 0$ if $(i,j) \in \mathcal{E}$, and zero otherwise; $\mathcal{N}(i) \triangleq \{j \in \mathcal{N} \mid (i,j) \in \mathcal{E}\}$ refers to the set of neighbors of node i; $\deg(i) \triangleq \sum_{j\in\mathcal{N}} a_{ij}$ will refer to the degree of each node $i \in \mathcal{N}$. Denote by Σ the N-dimensional probability simplex over $\mathcal{N}$, with $\text{Int}(\Sigma)$ denoting its interior, and let $\mathbf{P} \triangleq [P_{ij}]_{i,j\in\mathcal{N}}$ be the transition probability matrix of an ergodic, time-reversible Markov chain over $\mathcal{N}$, with its stationary distribution $\boldsymbol{\mu} \triangleq [\mu_i]_{i\in\mathcal{N}}$. Without loss of generality, we assume $P_{ij} > 0$ if and only if $a_{ij} > 0$. In this setup, we design *Self-Repellent Random Walks* (SRRWs) on general graphs[1] indexed by a tunable parameter $\alpha \geq 0$, all of which can sample from $\boldsymbol{\mu} \in \Sigma$, and then study their sampling 'efficiency' as a function of α (with $\alpha = 0$ being equivalent to the baseline Markov chain with transition kernel $\mathbf{P}$).

The SRRW transition kernel: Consider the Markov chain kernel (transition matrix) $\mathbf{K}[\mathbf{x}] \triangleq [K_{ij}[\mathbf{x}]]_{i,j\in\mathcal{N}}$, whose transition probabilities are mappings $K_{ij} : \Sigma \to [0,1]$, given by

$$K_{ij}[\mathbf{x}] \triangleq \frac{P_{ij} r_{\mu_j}(x_j)}{\sum_{k\in\mathcal{N}} P_{ik} r_{\mu_k}(x_k)}, \tag{1}$$

for any probability vector $\mathbf{x} \triangleq [x_i]_{i\in\mathcal{N}} \in \Sigma$. Here, $\{r_{\mu_i}\}_{i\in\mathcal{N}}$ is a family of positive functions $r_{\mu_i} : [0,1] \to \mathbb{R}_+$ parameterized by μ_i, with $r_{\mu_i}(x_i)$ *decreasing* in $x_i \in [0,1]$ and $r_{\mu_i}(\mu_i) = C$, for all $i \in \mathcal{N}$.[2] Transition probability kernels defined in this fashion, taking probability distributions as argument, are called *'nonlinear'* Markov kernels (Andrieu et al., 2007a; 2011), as opposed to classical Markov chains with kernels $\mathbf{P}$ that are often interpreted as linear operators – the transition probabilities at each step being independent of $\mathbf{x}$ (i.e., the case where $r_{\mu_i}(\cdot)$ is a constant function).

Stochastic processes utilizing nonlinear Markov kernels are called nonlinear Markov chains, and can be simulated/generated using *self-interacting Markov chains* (SIMCs) (see Del Moral & Miclo, 2004; 2006; Moral & Doucet, 2010). Let $\{X_n\}_{n\geq 0}$ be a random walker over $\mathcal{N}$, and let $\mathbf{x}_n$ be its *occupational measure* or *historical empirical distribution* up to time $n \geq 0$, written as

$$\mathbf{x}_n \triangleq \frac{1}{n+1} \sum_{k=0}^{n} \boldsymbol{\delta}_{X_k}, \tag{2}$$

where $\boldsymbol{\delta}_{X_k}$ is the delta measure whose X_k'th entry is one and the rest are zero, thus recording the position of the random walker at time $k \geq 0$. The process $\{X_n\}_{n\geq 0}$ becomes a SIMC if at each time step $n \geq 0$, the random walker makes transitions according to some nonlinear kernel $\mathbf{K}[\mathbf{x}_n]$, not necessarily as defined in (1). We say that the process $\{X_n\}_{n\geq 0}$ is a SRRW if it is a SIMC with $\mathbf{K}[\mathbf{x}]$ as in (1). We use the term *self-repellent* since at each time step, the transition probability to a node $j \in \mathcal{N}$ is proportional to $r_{\mu_j}([\mathbf{x}_n]_j)$ where $[\mathbf{x}_n]_j = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{X_k = j\}}$, and is thus a decreasing function of the visit count to $j \in \mathcal{N}$. In other words, the walker is less likely to move to a node that has been visited more often so far (thus self-repellent).

When $P_{ij} \propto a_{ij}$ in (1) for each $i \in \mathcal{N}$, the SRRW is a self-repellent version of the well-known *simple random walk* (SRW) procedure, with the target distribution being proportional to the degree of the nodes, that is, $\mu_i \propto \deg(i)$ for all $i \in \mathcal{N}$. Like most general MCMC procedures, the SRRW kernel can also be defined for *any given* sampling distribution $\boldsymbol{\mu} \in \text{Int}(\Sigma)$, for instance, by setting $\mathbf{P}$ to be the transition matrix of a Metropolis Hastings Random Walk (MHRW) with stationary distribution $\boldsymbol{\mu}$. For example if $\mu_i = 1/N$, that is $\boldsymbol{\mu} = \frac{1}{N}\mathbf{1}$ – the uniform distribution over the set of nodes $\mathcal{N}$, then we can choose $P_{ij} = \min\left\{\frac{1}{\deg(i)}, \frac{1}{\deg(j)}\right\}$ for all $(i,j) \in \mathcal{E}$, with $P_{ii} = 1 - \sum_{j\neq i} P_{ij}$. The matrix $\mathbf{P}$ defined in this manner is the MHRW kernel with the uniform distribution as its stationary measure, and is among the most commonly used kernels for unbiased graph sampling (Li et al., 2015) and distributed optimization (Sun et al., 2018). The elegance in the Metropolis Hastings algorithm and the key to its widespread adaptation lies in the fact that at each time step, the entries of $\boldsymbol{\mu}$ need only to be known for the neighbouring nodes of the random walkers (that is, only *local* information required), and only up to a constant multiple. This property ensures a robust, scalable implementation of the MHRW, since global constants are often unknown for large networks a priori.

Our SRRW construction begins with $r_{\mu_i}(\cdot)$ taking a polynomial form for all $i \in \mathcal{N}$, given by

$$r_{\mu_i}(x_i) \triangleq \left(\frac{x_i}{\mu_i}\right)^{-\alpha}, \quad \forall\, \alpha \geq 0, \tag{3}$$

where the parameter $\alpha \geq 0$ can be perceived as the *strength of the self-repellence mechanism* designed into the SRRW transition kernel. Similar to the MHRW transition kernel, only the local information regarding entries of $\boldsymbol{\mu}$ needs to be known at any given time step, and up to only a constant multiple. For convenience, we formalize this property as *scale-invariance (S.I.)*: an SRRW kernel posesses S.I. if for all $i, j \in \mathcal{N}$

(i) Computing $K_{ij}[\mathbf{x}]$ only requires knowing μ_k for $k \in \mathcal{N}(i)$, and only up to a constant multiple for any $i \in \mathcal{N}$;

[1] We consider graphs because they represent a generalization of (discrete) finite state spaces by imposing a communication (adjacency) matrix. The existence of an edge between two nodes (states) represents a non-zero probability of state transitions between the two nodes.

[2] As we shall see later, x_i will be directly proportional to the visit count to any node $\in \mathcal{N}$, since $vx \in \Sigma$ will be the empirical distribution of the self-repellent random walk.

(ii) $K_{ij}[C'\mathbf{x}] = K_{ij}[\mathbf{x}]$ for any constant $C' > 0$.

Indeed, we show in Appendix D that out of all possible forms for the functions $r_{\mu_i}(\cdot)$, only the polynomial form as in (3) possesses the S.I. property. Henceforth, we restrict ourselves to the polynomial form of $r_{\mu_i}(\cdot)$.

Our Contribution:

1. We show that given any MCMC kernel $\mathbf{P}$ which samples from a target distribution $\boldsymbol{\mu}$, the corresponding SRRW is asymptotically more *efficient* as a random walk based sampler. We do this by first showing that

$$\mathbf{x}_n \xrightarrow[n\to\infty]{\text{a.s.}} \boldsymbol{\mu}, \quad \forall\, \alpha \geq 0. \tag{4}$$

   We then provide second-order convergence results in the form of a central limit theorem (CLT); that is, we show there exists an *asymptotic co-variance matrix* $\mathbf{V}(\alpha) \in \mathbb{R}^{N\times N}$ parameterized by $\alpha \geq 0$, such that

$$\sqrt{n}(\mathbf{x}_n - \boldsymbol{\mu}) \xrightarrow[dist.]{n\to\infty} N(\mathbf{0}, \mathbf{V}(\alpha)). \tag{5}$$

   We obtain these results by first viewing the SRRW as a stochastic approximation (SA) algorithm with state-dependent noise (Harold J. Kushner, 1997; Fort, 2015), allowing us to form a connection between the stochastic process and a deterministic system of ordinary differential equations (ODEs). We establish global convergence results for this ODE system, which lay the foundations for proving almost sure convergence of the stochastic SRRW process.

2. For any $\alpha \geq 0$ we derive the *exact* form of $\mathbf{V}(\alpha)$ in terms of α and the spectrum (eigenvalues and eigenvectors) of $\mathbf{P}$. This allows us to show that kernels parameterized by larger α are asymptotically more *efficient* samplers; that is, they achieve smaller sampling variance. This is done by showing that the asymptotic covariance matrices follow a *Loewner ordering*:[3]

$$\mathbf{V}(\alpha_1) <_L \mathbf{V}(\alpha_2), \quad \forall\, \alpha_1 > \alpha_2 \geq 0. \tag{6}$$

   In other words, as long as the numerical/computational stability of the random walk implementation can be ensured, larger values of α are always more favourable in terms of achieving a smaller (asymptotic) sampling variance. We also derive an upper bound on the ratio of its sampling variance over that of the baseline Markov chain, and show that this upper bound *goes down to zero* as $\alpha \to \infty$ with speed $O(1/\alpha)$. This is surprising because asymptotically for large enough α, the SRRW, which is a stochastic process whose trajectories are constrained by 'walking' on the underlying communication matrix of the network, achieves smaller sampling variance than an *i.i.d.* sampler[4] whose variance is always a constant positive value.

3. We confirm our theoretical results by numerically observing the predicted asymptotic performance ordering of SRRW over a wide range of $\alpha \geq 0$ for the task of MCMC sampling over various graph topologies. To effectively handle potentially slower mixing of the SRRW process in the initial transient period for large α, we provide simulation results for an SRRW with *time-varying* α. Starting with smaller values of α which monotonically increase with time, we show that the empirically observed superior mixing of processes with smaller α can be combined with the theoretically proven asymptotic efficiency of SRRW with larger α, resulting in a far more efficient MCMC algorithm.

Related Works: While numerous version of self-repelling walks on countable state spaces have been studied in the past, we are primarily interested in random walk kernels which do not absolutely forbid (with probability one) transitions to previously visited nodes. The SRRW defined in our paper, which still allows transitions to past visited nodes with positive probability, falls into a class of 'weakly' self-avoiding random walks (Amit et al., 1983; Toth, 1995; Veto & Toth, 2008; Grassberger, 2017). The works *ibid.* study weakly self-avoiding random walks on topologies such as 1 or d-dimensional lattices, and provide theoretical results on properties such as recurrence properties, escape times from sets or average cover times. Thus, the analytical focus is not on convergence properties of some statistical attributes of the random walks to a particular target, making it difficult to utilize these existing results to design algorithms for learning and statistical inference on *general graphs*.

Works which do study convergence properties of empirical distributions of random walks with repellent dynamics consider processes where the repellence is between two (Chen, 2014) or more (Rosales et al., 2022) particles, and the results are again limited to random walks on complete graphs. Even thought the random walks therein are not self-repellent in our sense, the analysis techniques used are similar to ours, and are substantially influenced by the stochastic approximation framework in (Benaïm, 1997) for the analysis of vertex *reinforced* random walks (VRRW) (also see Benaïm & Tarres, 2011; Benaïm et al., 2012). As the name

[3] Matrices $\mathbf{A}, \mathbf{B}$ follow the Loewner ordering $\mathbf{A} \leq_L \mathbf{B}$ ($\mathbf{A} <_L \mathbf{B}$) if $\mathbf{B} - \mathbf{A}$ is positive semi-definite (positive definite).

[4] This corresponds to a sampler that can visit any node i with probability μ_i independent of its previous position at any given time. Clearly, in the graph setting, this requires the sampler to 'jump' to any other node by ignoring the underlying network structure altogether - something which random walkers on general graphs are not permitted to do.

might suggest, the nonlinear Markov kernels considered in this literature are self-attractive, that is, the random walker transitions to more frequently visited nodes with greater probability. The formulation is very general, and the attraction is captured by generic, monotonically *increasing* functions, which is in contrast to the monotonically *decreasing* $r_{\mu_i}(x_i)$ terms in our SRRW kernel as in (1). Even with such a general formulation, results showing convergence of the empirical measure to identifiable probability distributions are restricted to special cases such as complete graphs (Benaïm et al., 2012). Consequently, the analytical focus is shifted towards providing *localization* results, analyzing the property of the VRRW to get trapped inside subsets of the graph under consideration (Tarrès, 2004; Angel et al., 2014). Thus, even though the dynamics of VRRWs may look similar at first to the ones studied in our paper, the two types of random walks are essentially opposites of each other. The SRRW is specifically constructed for *sampling from any arbitrarily given target* distribution over any *general graph*, while the VRRW has found applications in fields such as ecology for understanding the behaviour of organisms like certain bacteria (Stevens & Othmer, 1997; Pemantle, 2007), and can be interpreted as *learning an unknown target* with potential applications to solving certain optimization problems (Avrachenkov et al., 2021) *over a special graph* such as a grid or a complete graph. Note that the SRRW kernel with $\alpha < 0$ in (3) is actually self-attractive, and therefore a VRRW. Our analytical approach covers such cases, and our first and second order convergence results actually hold for all α greater than a certain negative threshold value. For values of α smaller than this threshold (that is, the dynamics are more self-attractive in nature), we numerically observe non-convergence to the target distribution. This is analogous to the phase-transition behaviour often observed for VRRWs (Pemantle, 1988; Volkov, 2006; Akian et al., 2007) and puts our work in line with the broader literature of reinforced random walks.

As mentioned earlier, the SRRW described in this paper is a self-interacting Markov chain (SIMC), which can themselves be categorized under *adaptive* MCMC algorithms, with the kernel parameter being the empirical distribution of the process itself (Fort et al., 2011; 2014). While numerous works study such processes on continuous state spaces (Andrieu et al., 2007a;b; 2008; 2011), they focus on the construction of kernels employing *random jumps* to other (potentially distant) states depending on their historical visit counts, and are typically designed to tackle specific challenges such as sampling from multi-modal distributions and sequential Monte Carlo sampling, with performance improvements typically shown via simulation results. In practice, however, information regarding the shape of the target distribution such as its modality is unknown. These methods are also largely inapplicable for finite state spaces such as graphs where the underlying communication matrix determines the possible state transitions, therefore prohibiting jumps between non-neighbouring nodes.

Structure of the paper: In the rest of the paper, we first present some preliminaries and set up the SRRW procedure as a stochastic approximation in Section 2. In Section 3, we then analyze a mean-field ODE system whose asymptotic dynamics are closely related to our stochastic process; proving uniqueness of the target distribution $\boldsymbol{\mu} \in \text{Int}(\Sigma)$ as its fixed point and showing global convergence of its solutions to $\boldsymbol{\mu}$ with the help of a Lyapunov function. Our main results corresponding to our contributions as discussed earlier in the introduction are provided in Section 4, where we also discuss the application of our SRRW scheme to MCMC sampling. In Section 5 we provide numerical results supporting our theoretical findings along with additional tests that showcase the advantages of employing the SRRW with time-varying α, before concluding in Section 6.

2. Algorithmic setup

In this section, we first standardize some basic notations which will be used throughout the paper, and review preliminary concepts regarding time-reversible Markov chains and their spectrum. We then introduce the SRRW iteration as a stochastic approximation algorithm along with a deterministic system whose asymptotic dynamics are closely related to our stochastic iteration.

2.1. Preliminaries

Basic Notations: We use lower-case boldface letters to denote column vectors (e.g. $\mathbf{v} \triangleq [v_i]_{i\in\mathcal{N}} \in \mathbb{R}^N$), and upper-case boldface letters to denote matrices (e.g. $\mathbf{M} \triangleq [M_{ij}]_{i,j\in\mathcal{N}} \in \mathbb{R}^{N\times N}$). The matrix $\mathbf{I}$ always denotes the identity matrix (its dimension inferred by the context) and for any vector $\mathbf{v} \in \mathbb{R}^N$, we define $\mathbf{D}_{\mathbf{v}}$ as the diagonal matrix with v_i as its i-th diagonal entry. Throughout the paper, the terms $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ will stand for probability of an event and expectation of a random variable respectively, and we use $\|\cdot\|_p$ to refer to the L^p norm.

Time-Reversible Markov Chains: The following definitions and properties are classical in Markov chain literature and will be stated without proof. We point the readers to (Aldous & Fill, 2002, Chapter 3.4) for a more comprehensive review of the spectral properties stated below. A Markov chain with kernel $\mathbf{P} \in [0,1]^{N\times\mathbb{N}}$ is *time-reversible* if there exists a probability distribution $\boldsymbol{\mu} \in \Sigma$ such that the pair $(\mathbf{P}, \boldsymbol{\mu})$ solve the *detailed balance equation* (DBE), that is, $\mu_i P_{ij} = \mu_j P_{ji}$ for all $i, j \in \mathcal{N}$. Consequently, $\boldsymbol{\mu}$ is also the *stationary distribution* or *invariant measure* of $\mathbf{P}$, and solves $\boldsymbol{\mu}^T\mathbf{P} = \boldsymbol{\mu}^T$. A time-reversible Markov chain's kernel $\mathbf{P}$

has all eigenvalues on the real line. We denote by $(\lambda_i, \mathbf{u}_i)$ (by $(\lambda_i, \mathbf{v}_i)$) its left (right) eigenpair, where the eigenvalues are ordered such that $1 = \lambda_N > \lambda_{N-1} \geq \cdots \lambda_1 \geq -1$, with $\mathbf{u}_N = \boldsymbol{\mu}$ and $\mathbf{v}_N = \mathbf{1}$. Moreover, we can deduce that $\mathbf{u}_i = \mathbf{D}_{\boldsymbol{\mu}}\mathbf{v}_i$ and $\mathbf{u}_i^T\mathbf{v}_i = 0$ for all $i \in \mathcal{N}$.

2.2. SRRW iteration and the related mean-field ODE

Consider the SRRW on $\mathcal{G}(\mathcal{N}, \mathcal{E})$ as defined in Section 1, where for any ergodic and time-reversible $\mathbf{P}$ with corresponding stationary measure $\boldsymbol{\mu}$, and any $\alpha \geq 0$, the function $r_{\mu_i}(\cdot)$ in the non-linear kernel in (1) is of the polynomial form as in (3). In this setup, the transition kernel $\mathbf{K}[\mathbf{x}]$ is well-defined for all $\mathbf{x} \in \text{Int}(\Sigma)$, and we have the following result, whose proof is deferred to Appendix A.

Proposition 2.1. *For any $\mathbf{x} \in$ Int(Σ), there exists a unique stationary measure $\boldsymbol{\pi}(\mathbf{x}) \triangleq [\pi_i(\mathbf{x})]_{i\in\mathcal{N}} \in$ Int(Σ), where*

$$\pi_i(\mathbf{x}) \propto \sum_{j\in\mathcal{N}} \mu_i P_{ij} \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left(\frac{x_j}{\mu_j}\right)^{-\alpha}, \quad \forall\, i \in \mathcal{N}, \tag{7}$$

and the pair $(\mathbf{K}[\mathbf{x}], \boldsymbol{\pi}(\mathbf{x}))$ solves the DBE, that is, $\pi_i(\mathbf{x})K_{ij}[\mathbf{x}] = \pi_j(\mathbf{x})K_{ji}[\mathbf{x}]$ for all $i, j \in \mathcal{N}$.

Since the transition probability $K_{ij}[\mathbf{x}]$ for any node i to j is only well defined when $x_j > 0$ for each $j \in \mathcal{N}(i)$, we redefine $\mathbf{x}_n$ (originally defined in (2)) as $\mathbf{x}_n \triangleq \frac{1}{n+1}\left[\boldsymbol{\nu} + \sum_{k=1}^{n} \boldsymbol{\delta}_{X_k}\right]$ for some $\boldsymbol{\nu} \in \text{Int}(\Sigma)$. This redefined sequence of empirical measures $\{\mathbf{x}_n\}_{n\geq 0}$, which satisfies the recursion

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{1}{n+2}\left(\boldsymbol{\delta}_{X_{n+1}} - \mathbf{x}_n\right), \quad \forall\, n \geq 0, \tag{8}$$

starts from $\mathbf{x}_0 = \boldsymbol{\nu} \in \text{Int}(\Sigma)$, and satisfies $\mathbf{x}_n \in \text{Int}(\Sigma)$ which ensures that $\mathbf{K}[\mathbf{x}_n]$ is always well defined for every finite $n \geq 0$.

Remark 2.2. The S.I. property (ii) along with the form of $K_{ij}[\mathbf{x}]$ as in (1) means that knowing visit counts of each neighboring node of the current position is enough for computing the transition probabilities. This has benefits when the random walker is exploring the graph while performing MCMC sampling on-the-fly. For example, the random walker which associates one 'fake' visit to every newly discovered neighboring node automatically ensures that the SRRW iteration begins with $\boldsymbol{\nu} \propto \mathbf{1}$ (uniform distribution) without the need to know any *global* graph statistic. This 'fake' visits can also be weighted - assigning it proportional to the degree of the newly discovered neighboring node ensures that $\boldsymbol{\nu} = \mathbf{d}$, where $\mathbf{d} = [\deg(i)/\sum_{k\in\mathcal{N}} \deg(k)]_{i\in\mathcal{N}}$ is the degree proportional distribution of the graph. It is also important to note that our main results in Sections 4 are invariant in the choice of $\boldsymbol{\nu} \in \text{Int}(\Sigma)$. □

The existence and uniqueness of the stationary distribution $\boldsymbol{\pi}(\mathbf{x}_n)$ shown in Proposition 2.1 allows us to further decompose the recursion in (8) as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \gamma_{n+1}\left[\mathbf{h}(\mathbf{x}_n) + \boldsymbol{\epsilon}(X_{n+1}, \mathbf{x}_n)\right], \tag{9}$$

where $\gamma_n \triangleq \frac{1}{n+1}$, $\mathbf{h}(\mathbf{x}) \triangleq \boldsymbol{\pi}(\mathbf{x}) - \mathbf{x}$ and $\boldsymbol{\epsilon}(X, \mathbf{x}) \triangleq \boldsymbol{\delta}_X - \boldsymbol{\pi}(\mathbf{x})$. The recursion in (9) is an example of a stochastic approximation (SA) algorithm with controlled Markovian dynamic (see Benveniste et al., 2012; Andrieu et al., 2005; 2015; Fort et al., 2014) and decreasing step sizes $\{\gamma_n\}_{n\geq 0}$. For each $n \geq 0$, the random variable $\boldsymbol{\epsilon}(X_{n+1}, \mathbf{x}_n)$ captures the *noise* in the updates which, in this case, is driven by the stochastic process $\{X_n\}_{n\geq 0}$. SA algorithms have a strong connection to ordinary differential equations (ODEs) via its *mean-field*, which in our case is $\mathbf{h} : \text{Int}(\Sigma) \to \mathbb{R}^N$; SA iterates typically converge to ω-limit sets (fixed points and limit cycles) of the semi-flow induced by $\mathbf{h}(\cdot)$ as long as the noise terms $\{\boldsymbol{\epsilon}(X_n, \mathbf{x}_n)\}_{n\geq 0}$ are shown to have negligible contribution to the SA iterations asymptotically (Harold J. Kushner, 1997; Benaïm, 1999; Borkar, 2008). Studying the asymptotic properties of the *deterministic* ODE system given by

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{h}(\mathbf{x}(t)) = \boldsymbol{\pi}(\mathbf{x}(t)) - \mathbf{x}(t) \tag{10}$$

are therefore vital to understanding those of the *stochastic* iteration (9). For the SRRW kernel as in (1), we have $\mathbf{K}[\boldsymbol{\mu}] = \mathbf{P}$ for any $\alpha \geq 0$ since $r_{\mu_i}(\mu_i) = (\mu_i/\mu_i)^{-\alpha} = 1$ for all $i \in \mathcal{N}$; and from the balance equation $\boldsymbol{\pi}(\boldsymbol{\mu})^T\mathbf{P} = \boldsymbol{\pi}(\boldsymbol{\mu})^T$ we have $\boldsymbol{\pi}(\boldsymbol{\mu}) = \boldsymbol{\mu}$. Thus, the target measure $\boldsymbol{\mu}$ solves $\boldsymbol{\pi}(\mathbf{x}) = \mathbf{x}$, the fixed point equation of (10).

3. Global convergence of the mean-field ODE

In this section we analyze the asymptotic behaviour of the ODE, specifically the global asymptotic stability of the target distribution $\boldsymbol{\mu} \in \text{Int}(\Sigma)$. The ODE system in (10) is positively invariant in Int(Σ), that is, solutions of the system starting in Int(Σ) stay in the same set for all times $t \geq 0$. Indeed, one can check that as $x_i(t) \to 0$ for any $i \in \mathcal{N}$, the corresponding $\pi_i(\mathbf{x}) \to 1$, and the i-th entry $\frac{dx_i}{dt} = \pi_i(\mathbf{x}) - x_i$ of (10) will always be positive. With this, we first provide the following results concerning the fixed points of (10) in Int(Σ). The proofs of all results in this section are deferred to Appendix B.

Proposition 3.1. *For any $\alpha \geq 0$, the target stationary measure $\boldsymbol{\mu} \in$ Int(Σ) is the unique fixed point of ODE (10).*

Our next step is to show convergence of ODE (10) to its unique fixed point $\boldsymbol{\mu}$, which will be done via a *Lyapunov function*. For any $\alpha \geq 0$, consider the function $w : \text{Int}(\Sigma) \to \mathbb{R}$ given by

$$w(\mathbf{x}) = \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{N}} \mu_i P_{ij} \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left(\frac{x_j}{\mu_j}\right)^{-\alpha} \tag{11}$$

for all $\mathbf{x} \in \text{Int}(\Sigma)$. The following result shows that $w(\cdot)$ is a *strict* Lyapunov function for ODE (10).

Lemma 3.2. *The function $w : \text{Int}(\Sigma) \to \mathbb{R}$ satisfies $\nabla w(\mathbf{x})^T \mathbf{h}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \text{Int}(\Sigma)$, with equality only when $\mathbf{x} = \mu$. Moreover, $w(\mathbf{x}) \to \infty$ as $x_i \to 0$ for any $i \in \mathcal{N}$.*

The Lyapunov result in the form as stated above will be especially vital in Section 4 for proving the first and second order convergence results for the *stochastic* SRRW iteration (8). Our next result, concerning the global convergence of trajectories of (10), follows by application of the LaSalle invariance principle (see Vidyasagar, 1993, Chapter 5) to Lemma 3.2.

Theorem 3.3. *For any $\alpha \geq 0$, the unique fixed point μ of ODE (10) is globally asymptotically stable in Int(Σ); that is, for any initial $\mathbf{x}(0) \in$ Int(Σ), we have $\mathbf{x}(t) \to \mu$ as $t \to \infty$.*

For each $\alpha \geq 0$, let $\mathbf{J}(\alpha) \triangleq [J_{ij}(\alpha)]_{i,j \in \mathcal{N}}$ denote the Jacobian of $\mathbf{h}(\mathbf{x})$ evaluated at $\mathbf{x} = \boldsymbol{\mu}$, that is, $J_{ij}(\alpha) \triangleq \left.\frac{\partial h_i(\mathbf{x})}{\partial x_j}\right|_{\mathbf{x}=\boldsymbol{\mu}}$. We end this section with the following result regarding the form of $\mathbf{J}(\alpha)$ and its spectrum in terms of α, $\mathbf{P}$, and the spectrum of $\mathbf{P}$. This result will be employed to characterize the performance of SRRW in the next section.

Lemma 3.4. *For any $\alpha \geq 0$, we have*

$$\mathbf{J}(\alpha) = 2\alpha \boldsymbol{\mu} \mathbf{1}^T - \alpha \mathbf{P}^T - (\alpha+1)\mathbf{I}. \tag{12}$$

Furthermore, let ζ_i denote the i-th eigenvalue of $\mathbf{J}(\alpha)$. Then, $\zeta_N = -1$, and $\zeta_i = \alpha(-1-\lambda_i) - 1$ for all $i \in \{1, \cdots, N-1\}$, with $\mathbf{u}_i$ ($\mathbf{v}_i$) of $\mathbf{P}$ now being the right (left) eigenvectors corresponding to ζ_i for all $i \in \mathcal{N}$.

4. Main results - Convergence and co-variance ordering of SRRW

In this section, we provide our main results concerning the convergence properties of the SRRW iterate sequence $\{\mathbf{x}_n\}_{n \geq 0}$ for all $\alpha \geq 0$. First, we prove that starting from any initial $\mathbf{x}_0 \in \text{Int}(\Sigma)$, the SRRW iterate sequence $\{\mathbf{x}_n\}_{n \geq 0}$ satisfying (9) converges almost surely to the target distribution $\boldsymbol{\mu}$, and then provide a CLT where we characterize the exact form of the arising asymptotic co-variance matrix $\mathbf{V}(\alpha)$ as a function of $\alpha \geq 0$. We use this to prove our performance ordering as a Loewner ordering of asymptotic co-variance matrices, in the form of (6), as touched upon previously in Section 1. We also provide corollaries accompanying our main result for the special case of MCMC sampling, showing that the asymptotic sampling variance decreases in $\alpha \geq 0$ with speed $O(1/\alpha)$. All the following results are proved under the assumption (A1) as shown below, which we elaborate later in Remark 4.5.

(A1). For any $(\mathbf{x}_0, X_0) \in \text{Int}(\Sigma) \times \mathcal{N}$, the iterate sequence $\{\mathbf{x}_n\}_{n \geq 0}$ is $\mathbb{P}_{X_0, \mathbf{x}_0}$ - almost surely contained within a compact subset of Int(Σ).

Theorem 4.1 (Almost sure convergence to target $\boldsymbol{\mu}$)**.** *Under (A1), $\mathbf{x}_n$ converges $\mathbb{P}_{\mathbf{x}_0, X_0}$-almost surely to $\boldsymbol{\mu}$ as $n \to \infty$ for any initial $(\mathbf{x}_0, X_0) \in$ Int$(\Sigma) \times \mathcal{N}$ and any $\alpha \geq 0$.*

Theorem 4.2 (Central Limit Theorem)**.** *Under (A1),*

$$\sqrt{n}(\mathbf{x}_n - \boldsymbol{\mu}) \xrightarrow[dist.]{n \to \infty} N(\mathbf{0}, \mathbf{V}(\alpha)), \tag{13}$$

for any $\alpha \geq 0$, where $\mathbf{V}(\alpha) \in \mathbb{R}^{N \times N}$ is given by

$$\mathbf{V}(\alpha) = \sum_{i=1}^{N-1} \frac{1}{2\alpha(1+\lambda_i)+1} \cdot \frac{1+\lambda_i}{1-\lambda_i} \mathbf{u}_i \mathbf{u}_i^T. \tag{14}$$

The explicit form of $\mathbf{V}(\alpha)$ in Theorem 4.2 is derived by applying Lemma 3.4 in conjunction with a version of (Brémaud, 2020, Lemma 6.3.7) modified for our SRRW case. This allows us to fully characterize the co-variance matrix $\mathbf{V}(\alpha)$, and we have the following performance ordering result as a corollary.

Corollary 4.3 (Ordering of asymptotic covariance)**.** *For any $\alpha_1 > \alpha_2 > 0$, we have*

$$\mathbf{V}(\alpha_1) <_L \mathbf{V}(\alpha_2) <_L \mathbf{V}(0).$$

Corollary 4.3 states the asymptotic covariance matrix decrease monotonically (in terms of Loewner ordering) as α increases. Recall that the case $\alpha = 0$ coincides with a purely Markovian random walker with kernel $\mathbf{P}$, traversing the state space with no self interaction. Thus, the empirical distribution of any SRRW with $\alpha > 0$ approximates the target distribution with asymptotically smaller co-variance than the underlying base Markov chain $\mathbf{P}$. The CLT result in (13) also leads to the following corollary on the convergence of L^p norms of the scaled error term $\mathbf{x}_n - \boldsymbol{\mu}$.

Corollary 4.4. *for any $\alpha \geq 0$ and integer $p \in \mathbb{N}$, there exists a constant $C_{p,\alpha} > 0$ such that*

$$\sqrt{n}\mathbb{E}\left[\|\mathbf{x}_n - \mu\|_p\right] \xrightarrow{n \to \infty} C_{p,\alpha}. \tag{15}$$

Remark 4.5. (A1) assumes the *stability* of the SRRW iterates $\{\mathbf{x}_n\}_{n \geq 0}$, and is a commonplace in convergence analysis of SA algorithms. Proving (A1) for a given SA algorithm is highly non-trivial in most practical applications, and can be especially difficult for the case of SA with state-dependent Markovian noise, where the state space is an open subset of an $\mathbb{R}^N$ (Andrieu et al., 2015; Karmakar, 2020) - the case with our SRRW process. The issue of proving stability is often dealt with by introducing modifications to the original sequence, such as random restarts of the algorithm as the iterates exit a growing sequence of compact subsets of the open state space. In Appendix E, we use one such modification as a tool to provide intuition behind why (A1) is very likely to be satisfied, even though the formal proof remains an open problem. The critical logic behind this intuition revolves around the uniqueness of the target measure $\boldsymbol{\mu}$ and Lemma 3.2 regarding the Lyapunov function, and is further supported by our numerical simulations in Section 5. □

SRRW for MCMC sampling: We show that the performance improvement alluded by the form of $\mathbf{V}(\alpha)$ from Theorem 4.2 and Corollary 4.3 is realized when the SRRW is used for MCMC sampling. Consider a sampling agent following the SRRW process for an appropriate choice of $\mathbf{P}$ and $\boldsymbol{\mu}$, for any given $\alpha \geq 0$ on a general graph $\mathcal{G}(\mathcal{N}, \mathcal{E})$. At each time step $n > 0$, it records a sample according to its current location X_n, given by $g(X_n)$ for some scalar function $g : \mathcal{N} \to \mathbb{R}$, whose vectorized form we denote by $\mathbf{g} \triangleq [g(i)]_{i\in\mathcal{N}} \in \mathbb{R}^N$, and updates the MCMC *estimator* $\psi_n(g) \triangleq \frac{1}{n}\sum_{k=1}^{n} g(X_k)$ with the latest sample. Its goal is to estimate the quantity $\mathbf{g}^T\boldsymbol{\mu} = \mathbb{E}_{X\sim\mu}[g(X)]$. We now have the following:

Corollary 4.6. *For any scalar valued function $g : \mathcal{N} \to \mathbb{R}$ such that $\max_{i\in\mathcal{N}} |g(i)| < \infty$, and for any $\alpha \geq 0$, we have*

$$\psi_n(g) \xrightarrow[a.s.]{n\to\infty} \mathbf{g}^T\boldsymbol{\mu}, \tag{16}$$

$$\sqrt{n}(\psi_n(g) - \mathbf{g}^T\boldsymbol{\mu}) \xrightarrow[dist.]{n\to\infty} N(0, \mathbf{g}^T\mathbf{V}(\alpha)\mathbf{g}). \tag{17}$$

Corollary 4.7. *Under the same assumption of Corollary 4.6, we have*

$$\frac{\mathbf{g}^T\mathbf{V}(\alpha)\mathbf{g}}{\mathbf{g}^T\mathbf{V}(0)\mathbf{g}} \leq \mathbb{E}\left[\frac{1}{2\alpha(1+\Lambda)+1}\right], \tag{18}$$

where $\Lambda \in (-1, 1)$ is a random variable taking values λ_i, $i \in \{1, \cdots, N-1\}$ with probability proportional to $(\mathbf{g}^T\mathbf{u}_i)^2$.

Corollary 4.6 provides almost sure convergence and CLT for the SRRW-driven MCMC estimator $\psi_n(g)$, with its asymptotic variance in terms of $\mathbf{g}$ and $\mathbf{V}(\alpha)$. Corollary 4.7 then quantifies the improvement of the SRRW-driven sampler with $\alpha > 0$ over the baseline MCMC-driven by $\mathbf{P}$ ($\alpha = 0$ case). We see that reduction ratio goes down to zero as $\alpha \to \infty$ with speed $O(1/\alpha)$. Consider an *i.i.d.* random variable Y with distribution $\boldsymbol{\mu}$ for which the sampling variance is given by $\text{Var}(g(Y)) = (\mathbf{g}^2)^T\boldsymbol{\mu} - (\mathbf{g}^T\boldsymbol{\mu})^2 > 0$ for all non-trivial choices of function $g(\cdot)$ and μ.[5] Since $\mathbf{g}^T\mathbf{V}(\alpha)\mathbf{g} \to 0$ as $\alpha \to \infty$, there exists some $\hat{\alpha} > 0$ such that for each $\alpha > \hat{\alpha}$, the SRRW as an MCMC estimator outperforms even an i.i.d. sampler, asymptotically. While the value of $\hat{\alpha}$ depends on the function $g(\cdot)$, target measure $\boldsymbol{\mu}$ and the spectrum of the underlying base chain $\mathbf{P}$, we expect that this crossover would take place for moderate values of α, as seen from (18) with (at least) $O(1/\alpha)$ speed. This has the advantage of offsetting potentially undesirable transient behaviour resulting from employing very large α, as will also be elaborated in the next section while presenting our simulation results.

Before ending this section, we briefly show that for any choice of baseline Markov chain $\mathbf{P}$ with stationary distribution $\boldsymbol{\mu}$, one can construct an *unbiased* version of the estimator $\psi_n(g)$ which estimates the *uniform* average of the function $g(\cdot)$, given by $\frac{1}{N}\mathbf{g}^T\mathbf{1} = \mathbb{E}_{X\sim\text{unif.}}[g(X)]$. This procedure is virtually identical to the importance-reweighting typically done in MCMC applications (Lee et al., 2012), but now sampled by SRRW instead of a Markov chain. To be specific, consider a *weight function* $w : \mathcal{N} \to \mathbb{R}$ where $w(i) \propto \frac{1}{\mu_i}$ for all $i \in \mathcal{N}$, and let $h : \mathcal{N} \to \mathbb{R}$ be $h(i) \triangleq w(i)g(i)$ with $\mathbf{h} \triangleq [h(i)]_{i\in\mathcal{N}} \in \mathbb{R}^N$ being its vectorized form. Consider a new estimator, defined as

$$\hat{\psi}_n(g) \triangleq \frac{\psi_n(wg)}{\psi_n(w)} = \frac{\sum_{k=1}^{n} w(X_k)g(X_k)}{\sum_{k=1}^{n} w(X_k)}. \tag{19}$$

We then the following corollary, the proof of which follows by application of Slutsky's theorem (Ash & Doléans-Dade, 2000, pg. 332).

Corollary 4.8. *For the estimator $\hat{\psi}_n(g)$, results of Corollaries 4.6 and 4.7 hold with $\mathbf{g}$ and $\boldsymbol{\mu}$ replaced by $\mathbf{h}$ and $\frac{1}{N}\mathbf{1}$ respectively.*

5. Simulation results

We now present simulation results which support our theoretical findings. We simulate the SRRW process over the wiki-Vote graph (Leskovec & Krevl, 2014), which is an undirected, connected graph with 889 nodes and 2914 edge. We set $\mathbf{P}$, the *base* Markov chain, to be the Metropolis Hastings random walk (MHRW) - a commonly used Markov chain for unbiased graph sampling. Our target measure $\boldsymbol{\mu}$ is therefore the uniform distribution, that is, $\boldsymbol{\mu} = \frac{1}{N}\mathbf{1}$. We also measure the *average degree* of the graph by employing the estimator $\psi_n(g)$, where $g(i) = \deg(i)$ for each $i \in \mathcal{N}$, and $\mathbf{g}^T\boldsymbol{\mu} = \sum_{i\in\mathcal{N}} \text{deg(i)}/N$.

To track the convergence of the stochastic process, we measure the *Total Variation Distance* (TVD), evaluated at each time step as

$$\text{TVD}(\mathbf{x}_n, \boldsymbol{\mu}) = \frac{1}{2}\|\mathbf{x}_n - \boldsymbol{\mu}\|, \tag{20}$$

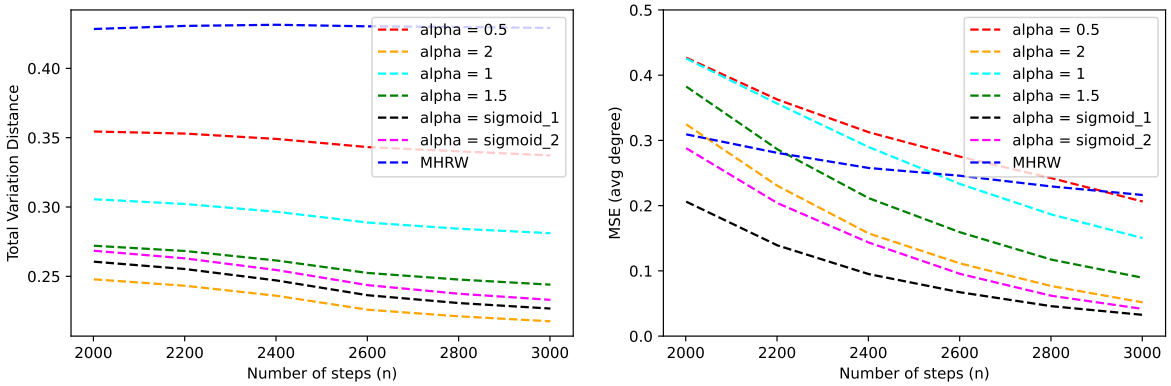
where $\mathbf{x}_n$ is the empirical distribution of the SRRW process. The TVD is then averaged over the total number of simulation runs. We also keep track of the *Mean Square Error* (MSE) of the estimator $\psi_n(g)$ in estimating the average degree, given by

$$\text{MSE}(\psi_n(g), \mathbf{g}^T\boldsymbol{\mu}) = \frac{1}{K}\sum_{k=1}^{K}(\psi_n^k(g) - \mathbf{g}^T\boldsymbol{\mu})^2 \tag{21}$$

where the *mean* is taken over the total number of simulation runs K, and $\psi_n^k(g)$ denotes the estimator for the k-th simulation run.

Our final simulation results are shown in Figures 1(a) and 1(b), where the curves are obtained as averages over $K =$

[5] Here, $\mathbf{g}^2$ is the vector with its i'th entry as $g(i)^2$.



(a) Convergence of $\mathbf{x}_n$ to the uniform distribution.

(b) Convergence of $\psi_n(g)$ to the ground truth $\mathbf{g}^T\mathbf{1}/N$.

Figure 1. Simulations of the SRRW process for values of $\alpha \in [0, 2]$, where $\alpha = 0$ corresponds to *MHRW* - the underlying Metropolis-Hastings *base* chain, with no self-repellence properties. The two *sigmoid* functions refer to the case where α is made to gradually increase over time, from 0 to 2. Sigmoid-1 is of the type $\frac{1}{0.5+e^{-n+0.25N}}$ while Sigmoid-2 is of the type $\frac{n}{100+0.5n}$, where $n \geq 0$ is the time parameter. Further tuning of the sigmoid functions may lead to empirically more efficient MCMC algorithms.

500 simulation runs. For both, TVD and MSE, we observe that larger values of α perform better asymptotically as predicted by our theory, significantly outperforming the case of $\alpha = 0$ (the sampler driven by base chain MHRW). We also observe for the MSE plot in Figure 1(b) that the case of time-varying α provides better performance for the task of estimating the average degree. While our theoretical framework only supports constant values of $\alpha \geq 0$, the numerically observed convergence and efficiency of the time-varying case opens up possibilities for construction of a more adaptive version of the SRRW algorithm, which can be a possible future direction.

6. Conclusion

In this paper, we introduced a *self-repellent random walk* which can be designed to sample from any target distribution $\boldsymbol{\mu} \in \text{Int}(\Sigma)$ via nonlinear interacting with its own occupational measure. We provide convergence results for the sequence of occupational measures by first analyzing a closely related deterministic process whose convergence is proved with the aid of a Lyapunov function, and then utilizing results from stochastic approximation theory to establish first and second order convergence results. We provide an explicit form for the asymptotic co-variance matrix arising out of the CLT result, which enables us to prove that the sampling variance decreases monotonically to zero as $\alpha \to \infty$, with speed at least $O(1/\alpha)$. Our results advocate for study into the design of nonlinear Markov chains for MCMC and other learning applications.

7. Acknowledgments and Disclosure of Funding

We thank the anonymous reviewers for their constructive comments. This research was largely conducted while Vishwaraj Doshi was with the Operations Research Graduate Program, North Carolina State University. This work was supported in part by National Science Foundation under Grant Nos. CNS-2007423, IIS-1910749, and CNS-1824518.

References

Akian, M., Gaubert, S., and Ninove, L. Multiple equilibria of nonhomogeneous markov chains and self-validating web rankings. *arXiv preprint arXiv:0712.0469*, 2007.

Aldous, D. and Fill, J. A. Reversible markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

Alon, N., Benjamini, I., Lubetzky, E., and Sodin, S. Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics*, 9(04):585–603, 2007.

Amit, D. J., Parisi, G., and Peliti, L. Asymptotic behavior of the "true" self-avoiding walk. *Phys. Rev. B*, 27:1635–1645, Feb 1983.

Andrieu, C. and Livingstone, S. Peskun–tierney ordering for markovian monte carlo: Beyond the reversible scenario. *The Annals of Statistics*, 49(4):1958–1981, 2021.

Andrieu, C., Moulines, É., and Priouret, P. Stability of stochastic approximation under verifiable conditions.

SIAM Journal on control and optimization, 44(1):283–312, 2005.

Andrieu, C., Jasra, A., Doucet, A., and Del Moral, P. Nonlinear markov chain monte carlo. *Esaim: Proceedings*, 19:79–84, 01 2007a.

Andrieu, C., Jasra, A., Doucet, A., and Del Moral, P. Convergence of the equi-energy sampler. In *ESAIM: Proceedings*, volume 19, pp. 1–5. EDP Sciences, 2007b.

Andrieu, C., Jasra, A., Doucet, A., and Del Moral, P. A note on convergence of the equi-energy sampler. *Stochastic Analysis and Applications*, 26(2):298–312, 2008.

Andrieu, C., Jasra, A., Doucet, A., and Moral, P. D. On nonlinear Markov chain Monte Carlo. *Bernoulli*, 17(3): 987 – 1014, 2011.

Andrieu, C., Tadić, V. B., and Vihola, M. On the stability of some controlled markov chains and its applications to stochastic approximation with markovian dynamic. *The Annals of Applied Probability*, 25(1):1–45, 2015.

Angel, O., Crawford, N., and Kozma, G. Localization for linearly edge reinforced random walks. *Duke Mathematical Journal*, 163(5):889–921, 2014.

Ash, R. B. and Doléans-Dade, C. A. *Probability and measure theory*. Academic press, 2000.

Avrachenkov, K. E., Borkar, V. S., Moharir, S., and Shah, S. M. Dynamic social learning under graph constraints. *IEEE Transactions on Control of Network Systems*, 9(3): 1435–1446, 2021.

Ball, J. Semiflows, Lyapunov functions and approach to equilibrium. University lecture, 2016. URL https://people.maths.ox.ac.uk/ball/Teaching/cdtsemiflows16.pdf.

Benaïm, M. Vertex-reinforced random walks and a conjecture of pemantle. *Annals of Probability*, 25(1):361–392, 01 1997.

Benaïm, M. *Dynamics of stochastic approximation algorithms*, pp. 1–68. Springer Berlin Heidelberg, 1999.

Benaïm, M. and Tarres, P. Dynamics of vertex-reinforced random walks. *The Annals of Probability*, 39(6):2178–2223, 2011.

Benaïm, M., Raimond, O., and Schapira, B. Strongly vertex-reinforced-random-walk on the complete graph. *arXiv preprint arXiv:1208.6375*, 2012.

Benveniste, A., Métivier, M., and Priouret, P. *Adaptive algorithms and stochastic approximations*, volume 22. Springer Science & Business Media, 2012.

Borkar, V. S. *Stochastic Approximation - A Dynamical Systems Viewpoint*. Hindustan Book Agency, 2008.

Brémaud, P. Markov chains gibbs fields, monte carlo simulation, and queues. 2020.

Chen, J. Two particles' repelling random walks on the complete graph. *Electronic Journal of Probability*, 19 (none):1 – 17, 2014.

Chen, T.-L. and Hwang, C.-R. Accelerating reversible markov chains. *Statistics & Probability Letters*, 83(9): 1956–1962, 2013.

Del Moral, P. and Miclo, L. On convergence of chains with occupational self-interactions. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 460(2041):325–346, 2004.

Del Moral, P. and Miclo, L. Self-interacting markov chains. *Stochastic Analysis and Applications*, 24:615–660, 07 2006.

Delyon, B. Stochastic approximation with decreasing gain: Convergence and asymptotic theory. Technical report, Université de Rennes, 2000.

Diaconis, P., Holmes, S., and Neal, R. M. Analysis of a nonreversible markov chain sampler. *Annals of Applied Probability*, pp. 726–752, 2000.

Fort, G. Central limit theorems for stochastic approximation with controlled markov chain dynamics. *ESAIM: Probability and Statistics*, 19:60–80, 2015.

Fort, G., Moulines, E., and Priouret, P. Convergence of adaptive and interacting markov chain monte carlo algorithms. *The Annals of Statistics*, 39(6):3262–3289, 2011.

Fort, G., Moulines, E., Priouret, P., and Vandekerkhove, P. A central limit theorem for adaptive and interacting markov chains. *Bernoulli*, 20(2):457–485, 2014.

Grassberger, P. Self-trapping self-repelling random walks. *Physical review letters*, 119(14):140601, 2017.

Harold J. Kushner, G. G. Y. *Stochastic Approximation Algorithms and Applications*. Springer, 1997.

Hu, J., Doshi, V., and Eun, D. Y. Efficiency ordering of stochastic gradient descent. *arXiv preprint arXiv:2209.07446*, 2022.

Karmakar, P. Stochastic approximation with markov noise: Analysis and applications in reinforcement learning. *arXiv preprint arXiv:2012.00805*, 2020.

Lee, C.-H., Xu, X., and Eun, D. Y. Beyond Random Walk and Metropolis-Hastings Samplers: Why You Should Not Backtrack for Unbiased Graph Sampling. In *Proceedings of the ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS'12, pp. 319–330, 2012.

Leskovec, J. and Krevl, A. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, June 2014.

Li, R.-H., Yu, J. X., Qin, L., Mao, R., and Jin, T. On random walk based graph sampling. In *2015 IEEE 31st international conference on data engineering*, pp. 927–938. IEEE, 2015.

Ma, Y.-A., Chen, T., Wu, L., and Fox, E. B. A unifying framework for devising efficient and irreversible mcmc samplers. *arXiv preprint arXiv:1608.05973*, 2016.

Moral, P. D. and Doucet, A. Interacting Markov chain Monte Carlo methods for solving nonlinear measure-valued equations. *The Annals of Applied Probability*, 20(2):593 – 639, 2010.

Neal, R. M. Improving asymptotic variance of mcmc estimators: Non-reversible chains are better. *arXiv preprint math/0407281*, 2004.

Pemantle, R. Phase transition in reinforced random walk and rwre on trees. *The Annals of Probability*, pp. 1229–1241, 1988.

Pemantle, R. A survey of random processes with reinforcement. *Probability surveys*, 4:1–79, 2007.

Rosales, R. A., Prado, F. P., and Pires, B. Vertex reinforced random walks with exponential interaction on complete graphs. *Stochastic Processes and their Applications*, 148: 353–379, 2022.

Slotine, J.-J. E. and Li, W. *Applied Nonlinear Control*. Prentice-Hall, 1991.

Stevens, A. and Othmer, H. G. Aggregation, blowup, and collapse: the abc's of taxis in reinforced random walks. *SIAM Journal on Applied Mathematics*, 57(4):1044–1081, 1997.

Sun, T., Sun, Y., and Yin, W. On markov chain gradient descent. *Advances in neural information processing systems*, 31, 2018.

Tarrès, P. Vertex-reinforced random walk on Z eventually gets stuck on five points. *The Annals of Probability*, 32 (3B):2650–2701, 2004.

Thin, A., Kotelevskii, N., Andrieu, C., Durmus, A., Moulines, E., and Panov, M. Nonreversible mcmc from conditional invertible transforms: a complete recipe with convergence guarantees. *arXiv preprint arXiv:2012.15550*, 2020.

Toth, B. The "True" Self-Avoiding Walk with Bond Repulsion on $\mathbb{Z}$: Limit Theorems. *The Annals of Probability*, 23(4):1523 – 1556, 1995.

Turitsyn, K. S., Chertkov, M., and Vucelja, M. Irreversible monte carlo algorithms for efficient sampling. *Physica D: Nonlinear Phenomena*, 240(4-5):410–414, 2011.

Van der Vaart, A. W. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Veto, B. and Toth, B. Self-repelling random walk with directed edges on Z. *Electronic Journal of Probability*, 13(none):1909 – 1926, 2008.

Vidyasagar, M. *Nonlinear Systems Analysis*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

Volkov, S. Phase transition in vertex-reinforced random walks on z with non-linear reinforcement. *J. Theoret. Probab*, 19(3):691–700, 2006.

A. Proof of results in Section 2

Proof of Proposition 2.1. It is enough to show that the form of $\boldsymbol{\pi}(\mathbf{x})$ specified in (7) satisfies the detailed balance equation. For any $i, j \in \mathcal{N}$, we have

$$\begin{aligned}
\pi_i(\mathbf{x})K_{ij}[\mathbf{x}] &\propto \left[\sum_{k\in\mathcal{N}} \mu_i P_{ik} \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left(\frac{x_k}{\mu_k}\right)^{-\alpha}\right] \frac{P_{ij}\left(\frac{x_j}{\mu_j}\right)^{-\alpha}}{\sum_{k\in\mathcal{N}} P_{ik}\left(\frac{x_k}{\mu_k}\right)^{-\alpha}} \times \frac{\mu_i}{\mu_i} \\
&= \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left[\sum_{k\in\mathcal{N}} \mu_i P_{ik} \left(\frac{x_k}{\mu_k}\right)^{-\alpha}\right] \frac{\mu_i P_{ij}\left(\frac{x_j}{\mu_j}\right)^{-\alpha}}{\sum_{k\in\mathcal{N}} \mu_i P_{ik}\left(\frac{x_k}{\mu_k}\right)^{-\alpha}} \\
&= \mu_i P_{ij} \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left(\frac{x_j}{\mu_j}\right)^{-\alpha}. \qquad (22)
\end{aligned}$$

Similarly, we can write

$$\pi_j(\mathbf{x})K_{ji}[\mathbf{x}] \propto \mu_j P_{ji} \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left(\frac{x_j}{\mu_j}\right)^{-\alpha}, \qquad (23)$$

and using the fact that $\mu_i P_{ij} = \mu_j P_{ji}$ due to the time-reversibility property of $\mathbf{P}$ with $\boldsymbol{\mu}$, (22) and (23) are equivalent, which completes the proof. □

B. Proof of results in Section 3

Proof of Proposition 3.1. We begin by using the form of $\boldsymbol{\pi}(\mathbf{x})$ derived in Proposition 2.1 to study the fixed point equation for ODE (10), which can be written for each $i \in \mathcal{N}$ as

$$x_i = \pi_i(\mathbf{x}) = \frac{1}{D}\sum_{j\in\mathcal{N}} \mu_i P_{ij} \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left(\frac{x_j}{\mu_j}\right)^{-\alpha},$$

where the normalizing constant D is given by $D = \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \mu_i P_{ij} \left(\frac{x_i}{\mu_i}\right)^{-\alpha} \left(\frac{x_j}{\mu_j}\right)^{-\alpha}$. Consider a change of variable where $y_i \triangleq x_i/\mu_i$ for all $i \in \mathcal{N}$, and rewrite the above equation to obtain

$$y_i = \frac{1}{D} y_i^{-\alpha} \sum_{j\in\mathcal{N}} P_{ij} y_j^{-\alpha}. \qquad (24)$$

It is enough to show that the only possible positive solutions to (24) are of the type $y_i = c > 0$ for all $i \in \mathcal{N}$. To check that it is indeed a solution, observe that

$$D = c^{-2\alpha} \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \mu_i P_{ij} = c^{-2\alpha} \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \mu_j P_{ji} = c^{-2\alpha} \sum_{j\in\mathcal{N}} \mu_j \sum_{i\in\mathcal{N}} P_{ji} = c^{-2\alpha},$$

where the second equality comes from the time-reversibility of $(\mu, \mathbf{P})$, which means that $\mu_i P_{ij} = \mu_j P_{ji}$ for all $i, j \in \mathcal{N}$, and the fourth equality comes from $\mathbf{P}$ being a stochastic matrix. Substituting $D = c^{-2\alpha}$ in (24) gives us

$$y_i = c = \frac{1}{c^{-\alpha}} c^{-\alpha} \sum_{j\in\mathcal{N}} P_{ij} c^{-\alpha} = 1,$$

implying $x_i = \mu_i$ for all $i \in \mathcal{N}$.

Suppose there exists another positive solution $\mathbf{y}'$ of (24) such that $\mathbf{y}' \neq c\mathbf{1}$ for some $c > 0$. Without loss of generality, we assume $y'_1 = \max_{i\in\mathcal{N}}\{y'_i\}$, $y'_N = \min_{i\in\mathcal{N}}\{y'_i\}$. Since $\mathbf{y}' \neq c\mathbf{1}$, we have $y'_1 > y'_N$, and from the fixed point equation (24) we have

$$y'_1 = \frac{1}{D}(y'_1)^{-\alpha} \sum_{j\in\mathcal{N}} P_{1j}(y'_j)^{-\alpha}, \qquad (25a)$$

$$y'_N = \frac{1}{D}(y'_N)^{-\alpha} \sum_{k\in\mathcal{N}} P_{Nk}(y'_k)^{-\alpha}. \tag{25b}$$

Dividing (25a) by (25b) and rearranging some terms gives us

$$\left(\frac{y'_1}{y'_N}\right)^{\alpha+1} = \frac{\sum_{j\in\mathcal{N}} P_{1j} y_j^{-\alpha}}{\sum_{k\in\mathcal{N}} P_{Nk} y_k^{-\alpha}}. \tag{26}$$

Since $\alpha \geq 0$, we have $(y'_1)^{-\alpha} \leq (y'_i)^{-\alpha} \leq (y'_N)^{-\alpha}$ for all $i \in \mathcal{N}$, and (26) has the upper bound

$$\left(\frac{y'_1}{y'_N}\right)^{\alpha+1} \leq \frac{(y'_N)^{-\alpha}\sum_{j\in\mathcal{N}} P_{1j}}{(y'_1)^{-\alpha}\sum_{k\in\mathcal{N}} P_{Nk}} = \left(\frac{y'_1}{y'_N}\right)^{\alpha},$$

which shows $y'_1/y'_N \leq 1$ and contradicts $y'_1 > y'_N$, which completes the proof.

We can additionally prove that that Proposition 3.1 holds true for all $\alpha \in (-0.5, 0)$ as well. When $\alpha < 0$, we have $(y'_N)^{-\alpha} \leq (y'_i)^{-\alpha} \leq (y'_1)^{-\alpha}$ for all $i \in \mathcal{N}$, and (26) has the upper bound

$$\left(\frac{y'_1}{y'_N}\right)^{\alpha+1} \leq \frac{(y'_1)^{-\alpha}\sum_{j\in\mathcal{N}} P_{1j}}{(y'_N)^{-\alpha}\sum_{k\in\mathcal{N}} P_{Nk}} = \left(\frac{y'_1}{y'_N}\right)^{-\alpha},$$

which leads to $(y'_1/y'_N)^{2\alpha+1} \leq 1$. If we additionally have $2\alpha + 1 > 0$, or equivalently $\alpha \in (-0.5, 0)$, we have $y'_1 \leq y'_N$, which contradicts $y'_1 > y'_N$, implying $x_i = \mu_i$ for all $i \in \mathcal{N}$ even when $\alpha \in (-0.5, 0)$ and the random walker is *self-attractive*. □

Proof of Lemma 3.2. We use the notation $\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}$ to denote a vector with the i'th entry being $\left[\frac{x_i}{\mu_i}\right]^{-\alpha}$, for any $\alpha \in \mathbb{R}$. For any vector $\mathbf{v} \in \mathbb{R}^N$, we will occasionally use the notation $[\mathbf{v}]_i$ to denote its i'th entry. Taking partial derivative of $w(\mathbf{x})$ with respect to x_i, we have

$$\frac{\partial w(\mathbf{x})}{\partial x_i} = \frac{-2\alpha}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i \tag{27}$$

Taking derivative of $w(\mathbf{x})$ along trajectories of the ODE system (10), we get

$$\begin{aligned}
\frac{d}{dt}w(\mathbf{x}) &= \sum_i \frac{\partial w(\mathbf{x})}{\partial x_i}\frac{dx_i}{dt} = \nabla w(\mathbf{x})^T \mathbf{h}(vx) = -2\alpha \sum_{i\in\mathcal{N}} \left(\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right)\cdot(\pi_i(\mathbf{x}) - x_i) \\
&= -2\alpha \sum_{i\in\mathcal{N}} \left(\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right)\cdot\left(\frac{1}{w(\mathbf{x})}\left[\frac{x_i}{\mu_i}\right]^{-\alpha}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i - x_i\right) \\
&= \frac{-2\alpha}{w(\mathbf{x})} \sum_{i\in\mathcal{N}} \left(\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right)\cdot\left(\left[\frac{x_i}{\mu_i}\right]^{-\alpha}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i - w(\mathbf{x})x_i\right) \\
&= \frac{-2\alpha}{w(\mathbf{x})} \sum_{i\in\mathcal{N}} B_i(\mathbf{x}) + C_i(\mathbf{x}),
\end{aligned} \tag{28}$$

where

$$B_i(\mathbf{x}) = \left(\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right)\cdot\left(\left[\frac{x_i}{\mu_i}\right]^{-\alpha}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right),$$

and

$$C_i(\mathbf{x}) = \left(\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right)\cdot(w(\mathbf{x})x_i).$$

Define a random variable $Z(\mathbf{x})$ which takes values $Z_i(\mathbf{x}) = \frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i$ with probability x_i. Then, we can see that

$$B_i(\mathbf{x}) = \left(\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right)\cdot\left(\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\right)\cdot x_i = Z_i(\mathbf{x})^2 x_i,$$

and as a result,

$$\sum_{i\in\mathcal{N}} B_i(\mathbf{x}) = \sum_{i\in\mathcal{N}} Z_i(\mathbf{x})^2 x_i = \mathbb{E}[Z(\mathbf{x})^2]. \tag{29}$$

We can similarly write $\sum_{i\in\mathcal{N}} C_i(\mathbf{x})$ as

$$\begin{aligned}\sum_{i\in\mathcal{N}} C_i(\mathbf{x}) &= -w(\mathbf{x})\sum_{i\in\mathcal{N}}\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\cdot x_i\\ &= -\left(\sum_{k\in\mathcal{N}}\frac{1}{\mu_k}\left[\frac{x_k}{\mu_k}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_k\cdot x_k\right)\left(\sum_{i\in\mathcal{N}}\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i\cdot x_i\right)\\ &= -\left(\sum_{k\in\mathcal{N}} Z_k(\mathbf{x})x_k\right)\left(\sum_{i\in\mathcal{N}} Z_i(\mathbf{x})x_i\right) = \mathbb{E}[Z(\mathbf{x})]^2.\end{aligned} \tag{30}$$

Substituting (29) and (30) in (28) gives us

$$\frac{d}{dt}w(\mathbf{x}) = \frac{-2\alpha}{w(\mathbf{x})}\sum_{i\in\mathcal{N}} B_i(\mathbf{x}) + C_i(\mathbf{x}) = \frac{-2\alpha}{w(\mathbf{x})}\left(\mathbb{E}[Z(\mathbf{x})^2] - \mathbb{E}[Z(\mathbf{x})]^2\right) = \frac{-2\alpha}{w(\mathbf{x})}\mathrm{Var}[Z(\mathbf{x})] \le 0. \tag{31}$$

To show that the equality is only achieved at the fixed point, all we need to show is that $\mathrm{Var}[Z(\mathbf{x})] = 0 \iff \mathbf{x} = \boldsymbol{\pi}(\mathbf{x})$. The Variance term is zero if and only if $Z_i = Z_j$ for all $i, j \in \mathcal{N}$, that is

$$\frac{1}{\mu_i}\left[\frac{x_i}{\mu_i}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_i = \frac{1}{\mu_i}\left[\frac{x_j}{\mu_j}\right]^{-\alpha-1}\left[\mathbf{D}_{\boldsymbol{\mu}}\mathbf{P}\left[\frac{\mathbf{x}}{\boldsymbol{\mu}}\right]^{-\alpha}\right]_j \iff \frac{\pi_i(\mathbf{x})}{x_i} = \frac{\pi_j(\mathbf{x})}{x_j} \tag{32}$$

for all $i, j \in \mathcal{N}$, where the equality on the left hand size comes by rewriting the form for Z_i as $w(\mathbf{x})\pi_i(\mathbf{x})/x_i$. Equation (32) is true if and only if $\pi_i(\mathbf{x}) = x_i$ for all $i \in \mathcal{N}$, which completes the proof. □

Proof of Theorem 3.3. The trajectories $\{\mathbf{x}(t)\}_{t\ge 0}$ starting from any $\mathbf{x}(0) \in \mathrm{Int}(\Sigma)$ are relatively compact and bounded, since the flow of the ODE system (10) leaves the probability simplex positively invariant. Then, the global asymptotic stability of $\boldsymbol{\mu} \in \mathrm{Int}(\Sigma)$ follows by application of Proposition 3.1, Lemma 3.2, and the LaSalle Invariance principle (Theorem 3.1 in (Ball, 2016), Chapter 5 in (Vidyasagar, 1993), Chapter 3 in (Slotine & Li, 1991)). □

Proof of Lemma 3.4. Consider $f_i(\mathbf{x})$ defined as

$$f_i(\mathbf{x}) \triangleq \sum_{j\in\mathcal{N}} \mu_i P_{ij}\left(\frac{x_i}{\mu_i}\right)^{-\alpha}\left(\frac{x_j}{\mu_j}\right)^{-\alpha}$$

for all $i \in \mathcal{N}$, $\mathbf{x} \in \mathrm{Int}(\Sigma)$, and let $g(\mathbf{x}) \triangleq \sum_{k\in\mathcal{N}} h_k(\mathbf{x})$. Then, for $h_i(\mathbf{x}) = \pi_i(\mathbf{x}) - x_i$ (the i'th entry of $\mathbf{h}(\mathbf{x})$ in (10)), we have

$$h_i(\mathbf{x}) = \frac{f_i(\mathbf{x})}{g(\mathbf{x})} - x_i, \tag{33}$$

for all $i \in \mathcal{N}$ and $\mathbf{x} \in \mathrm{Int}(\Sigma)$, and its partial derivatives follow

$$\frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}_j} = \frac{g(\mathbf{x})\frac{\partial f_i(\mathbf{x})}{\partial x_j} - f_i(\mathbf{x})\frac{\partial g(\mathbf{x})}{\partial x_j}}{g(\mathbf{x})^2} - \mathbb{1}_{\{i=j\}}. \tag{34}$$

To deduce $\frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}_j}\big|_{\mathbf{x}=\boldsymbol{\mu}}$, we evaluate each quantity of the above equation evaluated at $\mathbf{x} = \boldsymbol{\mu}$, and then substitute them back in the above expression. We have

$$
\begin{aligned}
f_i(\boldsymbol{\mu}) &= \sum_{j\in\mathcal{N}} \mu_i P_{ij} = \mu_i \sum_{j\in\mathcal{N}} P_{ij} = \mu_i, && \forall\, i \in \mathcal{N},\\
g(\boldsymbol{\mu}) &= \sum_{i\in\mathcal{N}} f_i(\boldsymbol{\mu}) = \sum_{i\in\mathcal{N}} \mu_i = 1,\\
\frac{\partial f_i(\mathbf{x})}{\partial \mathbf{x}_i}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} &= -\frac{\alpha}{x_i}\sum_{j\in\mathcal{N}} \mu_i P_{ij}\left(\frac{x_i}{\mu_i}\right)^{-\alpha}\left(\frac{x_j}{\mu_j}\right)^{-\alpha} - \frac{\alpha}{x_i}\mu_i P_{ii}\left(\frac{x_i}{\mu_i}\right)^{-2\alpha}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} = -\alpha - \alpha P_{ii}, && \forall\, i \in \mathcal{N},\\
\frac{\partial f_i(\mathbf{x})}{\partial \mathbf{x}_j}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} &= -\frac{\alpha}{x_j}\mu_i P_{ij}\left(\frac{x_i}{\mu_i}\right)^{-\alpha}\left(\frac{x_j}{\mu_j}\right)^{-\alpha}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} = -\alpha\frac{\mu_i P_{ij}}{\mu_j} = -\alpha P_{ji}, && \forall\, i \neq j \in \mathcal{N},\\
\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}_i}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} &= -\frac{2\alpha}{x_i}\sum_{j\in\mathcal{N}} \mu_i P_{ij}\left(\frac{x_i}{\mu_i}\right)^{-\alpha}\left(\frac{x_j}{\mu_j}\right)^{-\alpha}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} = -2\alpha, && \forall\, i \in \mathcal{N}.
\end{aligned}
$$

Substituting the above expressions in (34) and simplifying it yields

$$
\begin{aligned}
\frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}_i}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} &= 2\alpha\mu_i - \alpha P_{ii} - \alpha - 1, && \forall\, i \in \mathcal{N},\\
\frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}_j}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} &= 2\alpha\mu_i - \alpha P_{ji}, && \forall\, i \neq j \in \mathcal{N},
\end{aligned}
$$

and by rewriting the above in matrix form, we get

$$
\mathbf{J}(\alpha) = 2\alpha\boldsymbol{\mu}\mathbf{1}^T - \alpha\mathbf{P}^T - (\alpha+1)\mathbf{I} \tag{35}
$$

which is the same as (12).

We now prove the eigenvalue result in Lemma 3.4. For each μ_i for $i \in \{1, \cdots, N\}$, we have

$$
\mathbf{J}(\alpha)^T\mathbf{u}_i = 2\alpha\boldsymbol{\mu}\mathbf{1}^T\mathbf{u}_i - \alpha\mathbf{P}^T\mathbf{u}_i - (\alpha+1)\mathbf{u}_i.
$$

When $i = N$, that is $\mathbf{u}_i = \mathbf{u}_N = \boldsymbol{\mu}$, then we have $\boldsymbol{\mu}\mathbf{1}^T\mathbf{u}_N = \mathbf{u}_N\mathbf{1}^T\boldsymbol{\mu} = \boldsymbol{\mu} = \mathbf{u}_N$, and $\mathbf{P}^T\mathbf{u}_N = \mathbf{u}_N$, and so the above equation becomes

$$
\mathbf{J}(\alpha)^T\mathbf{u}_N = (2\alpha - \alpha - \alpha - 1)\mathbf{u}_N = (-1)\mathbf{u}_N.
$$

When $i \neq N$, we have $\boldsymbol{\mu}\mathbf{1}^T\mathbf{u}_i = \boldsymbol{\mu}\mathbf{v}_N^T\mathbf{u}_i = 0$, and $\mathbf{P}^T\mathbf{u}_i = \lambda_i\mathbf{u}_i$, and we similarly have

$$
\mathbf{J}(\alpha)^T\mathbf{u}_i = (0 - \alpha\lambda_i - \alpha - 1)\mathbf{u}_i = (\alpha(-1-\lambda_i) - 1)\mathbf{u}_i.
$$

Similar steps follow when we start from $\mathbf{v}_i^T\mathbf{J}(\alpha)^T$ instead, and $\zeta_N = -1$, $\zeta_i = (\alpha(-1-\lambda_i)-1)$ for all $i \in \{1, \cdots, N-1\}$ are the eigenvalues of $\mathbf{J}(\alpha)$ with $\mathbf{u}_i$ and $\mathbf{v}_i$ being the corresponding left and right eigenvectors. Since $(-\lambda_i - 1) < 0$ for all $i \in \{1, \cdots, N-1\}$, ζ_i's follow the same ordering as λ_i's and this completes the proof. □

C. Proof of results in Section 4

Before providing the proofs of our main results, we reproduce some key assumptions from (Delyon, 2000) required to apply Theorems 15 and 25 therein, which are the almost sure convergence and the CLT result respectively.

(A) $\mathbf{h}$ is a continuous vector field on $\mathcal{O} \subset \mathbb{R}^d$; there exists a non-negative C^1 function w such that

- $\nabla w(\mathbf{x})^T\mathbf{h}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathcal{O}$;
- the set $S \triangleq \{\mathbf{x} \mid \nabla w(\mathbf{x})^T\mathbf{h}(\mathbf{x}) = 0\}$ is such that $w(S)$ has an empty interior.

(B) $\mathbf{h}$ is a continuous vector field on $\mathcal{O} \subset \mathbb{R}^d$; there exists a non-negative C^1 function w and a compact set $\mathcal{K} \subset \mathcal{O}$ such that

- $w(\mathbf{x}) \to \infty$ if $\mathbf{x} \to \partial\mathcal{O}$ or $\|\mathbf{x}\| \to \infty$;
- $\nabla w(\mathbf{x})^T \mathbf{h}(\mathbf{x}) < 0$ if $x \notin \mathcal{K}$.

(C) The general SA iteration given by $\mathbf{x}_{n+1} = \mathbf{x}_n + \gamma_{n+1}[\mathbf{h}(\mathbf{x}_n) + \eta_{n+1}]$ is *A-stable* (see Delyon, 2000, Definition 1) and $\mathbf{x}_n$ converges to some limit $\mathbf{x}^*$. $\mathbf{h}$ is C^1 in some neighborhood of $\mathbf{x}^*$ with first derivatives Lipschitz, and the Jacobian matrix of $\mathbf{h}$ evaluated at $\mathbf{x}^*$ has all its eigenvalues with negative real part.

(D) The step size is decreasing and satisfies either

$$\begin{cases} \frac{1}{\gamma_n} - \frac{1}{\gamma_{n-1}} \to 0; \\ \gamma_n n \to 1. \end{cases}$$

(MS) (Translated to the non-linear Markov chain setting) For every $\mathbf{x}$, there exits a solution $\mathbf{Q}(\mathbf{x})$ to the *Poisson equation*

$$(\mathbf{I} - \mathbf{K}[\mathbf{x}])\mathbf{Q}[\mathbf{x}] = \mathbf{I} - \mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T$$

has a solution; for any compact $\mathcal{C} \subset \mathcal{O}$,

$$\sup_{\mathbf{x}\in\mathcal{K}, i\in\mathcal{N}} \|\mathbf{Q}[\mathbf{x}]^T \mathbf{e}_i\|_2 + \|(\mathbf{I} - \mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T)^T \mathbf{e}_i\|_2 < \infty$$

and there exists a continuous function $\phi_{\mathcal{C}}$, $\phi_{\mathcal{C}}(0) = 0$, such that for any $\mathbf{x}, \mathbf{x}' \in \mathcal{C}$,

$$\sup_{i\in\mathcal{N}} \| [\mathbf{K}[\mathbf{x}]\mathbf{Q}[\mathbf{x}]]_{\cdot,i} - [\mathbf{K}[\mathbf{y}]\mathbf{Q}[\mathbf{y}]]_{\cdot,i} \|_2 \le \phi_{\mathcal{C}}(\|\mathbf{x} - \mathbf{y}\|_2).$$

Proof of Theorem 4.1. As mentioned in Section 2, the SRRW iteration (9) is a stochastic approximation algorithm with controlled Markovian input, with its step size sequence given by $\gamma_n = \frac{1}{n+1}$. To prove the almost sure convergence, we show that assumptions (A), (B) and (MS) in (Delyon, 2000) are satisfied, and then, under (A1), apply Theorem 15 therein. As a result of Proposition 3.1 and Lemma 3.2, the set of fixed points, which is the singleton $\{\mu\} \subset \text{Int}(\Sigma)$, and the Lyapunov function $w : \text{Int}(\Sigma) \to [0, \infty)$ as defined in (11) satisfy assumptions (A) and (B) in (Delyon, 2000).

Since $\mathbf{K}[\mathbf{x}]$ is irreducible for all $\mathbf{x} \in \text{Int}(\Sigma)$, the semigroup $\{e^{t(\mathbf{K}[\mathbf{x}]-\mathbf{I})}\}_{t\ge 0}$ of the related CTMC kernel $\mathbf{K}[\mathbf{x}] - \mathbf{I}$ converges exponentially towards $\mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T$ (geometric ergodicity). Thus, for all $\mathbf{x} \in \text{Int}(\Sigma)$, or equivalently for all $\mathbf{x} \in \mathcal{K}$ for any compact $\mathcal{K} \subset \text{Int}(\Sigma)$, the matrix

$$\mathbf{Q}[\mathbf{x}] = \int_0^\infty \left(e^{t(\mathbf{K}[\mathbf{x}]-\mathbf{I})} - \mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T \right) dt$$

is well defined. Moreover, it solves the Poisson equation, that is

$$\begin{aligned}(\mathbf{I} - \mathbf{K}[\mathbf{x}])\mathbf{Q}[\mathbf{x}] &= \int_0^\infty \left((\mathbf{I} - \mathbf{K}[\mathbf{x}])e^{-t(\mathbf{I}-\mathbf{K}[\mathbf{x}])} - (\mathbf{I} - \mathbf{K}[\mathbf{x}])\mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T \right) dt \\ &= \int_0^\infty (\mathbf{I} - \mathbf{K}[\mathbf{x}])e^{-t(\mathbf{I}-\mathbf{K}[\mathbf{x}])} dt = \mathbf{I} - \mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T = \mathbf{I} - e^{t(\mathbf{K}[\mathbf{x}]-\mathbf{I})}\Big|_0^\infty = \mathbf{I} - \mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T\end{aligned}$$

where the last inequality is because $-e^{t(\mathbf{K}[\mathbf{x}]-\mathbf{I})}$ the semi-group operator of an ergodic CTMC. The solution of the Poisson equation $\mathbf{Q}[\mathbf{x}]$, as well as the state dependent update (matrix) $\mathbf{I} - \mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T$, have bounded entries for all $\mathbf{x} \in \mathcal{K} \subset \text{Int}(\Sigma)$, which implies that

$$\sup_{\mathbf{x}\in\mathcal{K}, i\in\mathcal{N}} \|\mathbf{Q}[\mathbf{x}]^T \mathbf{e}_i\|_2 + \|(\mathbf{I} - \mathbf{1}\boldsymbol{\pi}(\mathbf{x})^T)^T \mathbf{e}_i\|_2 = \|\mathbf{Q}[\mathbf{x}]_{\cdot,i}\|_2 + \|\boldsymbol{\delta}_i - \boldsymbol{\pi}(\mathbf{x})\|_2 < \infty.$$

Moreover, since $\mathbf{K}[\mathbf{x}]$ and $\mathbf{Q}[\mathbf{x}]$ are continuous, they are also Lipschitz in $\mathcal{K}$. Thus fore each $\mathcal{K}$, there exists a constant $C_{\mathcal{K}}$ such that for any $\mathbf{x}, \mathbf{y} \in \mathcal{K}$,

$$\sup_{i\in\mathcal{N}} \| [\mathbf{K}[\mathbf{x}]\mathbf{Q}[\mathbf{x}]]_{\cdot,i} - [\mathbf{K}[\mathbf{y}]\mathbf{Q}[\mathbf{y}]]_{\cdot,i} \|_2 \le C_{\mathcal{K}} \|\mathbf{x} - \mathbf{y}\|_2.$$

With this, we satisfy (MS) in (Delyon, 2000). □

Proof of Theorem 4.2. It suffices to show that assumptions (C), (D) and (MS) in (Delyon, 2000) are satisfied and apply Theorem 25 therein, and then analyzing the form of the resulting co-variance matrix to prove our CLT result. Since we already showed that (MS) is satisfied while proving Theorem 4.1, we focus on proving (C) and (D).

Our choice of step size, $\gamma_n = \frac{1}{n+2}$ satisfies $\sum_{n\in\mathbb{N}_0}\gamma_n = \infty$ and $\sum_{n\in\mathbb{N}_0}\gamma_n^2 < \infty$. Besides,

$$\gamma_n - \gamma_{n+1} = \frac{1}{n+2} - \frac{1}{n+3} = \frac{(n+3)-(n+2)}{(n+2)(n+3)} \leq \frac{1}{(n+2)^2}.$$

Then, we have $\sum_{n\in\mathbb{N}_0}|\gamma_n - \gamma_{n+1}| \leq \sum_{n\in\mathbb{N}_0} 1/(n+2)^2 < \infty$, and (D) in (Delyon, 2000) is satisfied.

Since our mean field $F(\mathbf{x})$ is differentiable everywhere in $\text{Int}(\Sigma)$ and therefore continuous, it is Lipschitz for all compact subsets $\mathcal{K} \subset \text{Int}(\Sigma)$, and thus also Lipschitz over some neighborhood of $\mathbf{d} \in \text{Int}(\Sigma)$. Moreover linear stability of $\mathbf{d}$ follows from the global stability of shown in Theorem 3.3, and all eigenvalues of $\mathbf{J_h}(\mathbf{d})$ have negative real parts. This ensures that (C) in (Delyon, 2000) is satisfied.

This completes the proof of (13), and we will now show (14).

To obtain the closed form of $\mathbf{U}$, we first provide Lemma 6.3.7 in (Brémaud, 2020)), but re-written for vector-valued functions instead.

Lemma C.1 (Lemma 6.3.7 in (Brémaud, 2020))**.** *Let $\{X_k\}_{k\geq 0}$ be an ergodic Markov chain (reversible) with finite state space $[n]$, transition probability matrix $\mathbf{P}$ and stationary distribution $\boldsymbol{\mu}$. For any function $\mathbf{f} : [n] \to \mathbb{R}^d$, we have*

$$U(\mathbf{f}) = 2\mathbf{F}^T diag(\boldsymbol{\pi})\mathbf{Z}\mathbf{F} - \mathbf{F}^T diag(\boldsymbol{\pi})\mathbf{F} - \mathbf{F}^T\boldsymbol{\mu}\boldsymbol{\mu}^T\mathbf{F}, \tag{36}$$

where $U(\mathbf{f})$ is the asymptotic covariance matrix for function $\mathbf{f}$, matrix $\mathbf{F}$ is given by $\mathbf{F} \triangleq [\mathbf{f}(1), \cdots, \mathbf{f}(n)]^T$, and $\mathbf{Z} \triangleq (\mathbf{I} - \mathbf{P} + \mathbf{1}\boldsymbol{\mu}^T)^{-1}$. Moreover, since $\mathbf{P}$ is reversible, we have (equation 6.34 in (Brémaud, 2020))

$$U(\mathbf{f}) = \sum_{k=1}^{n-1} \frac{1+\lambda_k}{1-\lambda_k}\mathbf{F}^T\mathbf{u}_k\mathbf{u}_k^T\mathbf{F}, \tag{37}$$

where $\mathbf{u}_k$ are the left eigenvectors of $\mathbf{P}$ with $\mathbf{u}_n = \boldsymbol{\mu}$.

For our update rule (10), we have $\mathbf{F} = \mathbf{I} - \mathbf{1}\boldsymbol{\mu}^T$.[6] The asymptotic covariance matrix $\mathbf{U}$ is the same for all $\alpha \geq 0$, since is given with respect to $\mathbf{F}$ and the transition kernel $\mathbf{P} = \mathbf{K}[\boldsymbol{\mu}]$. Thus, we can write down $\mathbf{U}$ as

$$\mathbf{U} = \sum_{k=1}^{N-1} \frac{1+\lambda_k}{1-\lambda_k}(\mathbf{I} - \boldsymbol{\mu}\mathbf{1}^T)\mathbf{u}_k\mathbf{u}_k^T(\mathbf{I} - \mathbf{1}\boldsymbol{\mu}^T) = \sum_{k=1}^{N-1} \frac{1+\lambda_k}{1-\lambda_k}\mathbf{u}_k\mathbf{u}_k^T \tag{38}$$

where the last equality is because $\mathbf{u}_N = \boldsymbol{\mu}$ and $\mathbf{v}_N = \mathbf{1}$, and since $\mathbf{u}_i^T\mathbf{v}_j = 0$ for all $i \neq j$.

Lemma 3.4 allows us to write down the spectral decomposition of $e^{t(\mathbf{J}(\alpha)+\mathbf{I}/2)}$ as

$$e^{t(\mathbf{J}(\alpha)+\mathbf{I}/2)} = \sum_{i\in\mathcal{N}} e^{t(\zeta_i+1/2)}\mathbf{u}_i\mathbf{v}_i^T = e^{-1/2}\mathbf{u}_N\mathbf{v}_N^T + \sum_{i=1}^{N-1} e^{t(\alpha(-1-\lambda_i)-1/2)}\mathbf{u}_i\mathbf{v}_i^T. \tag{39}$$

Substituting (38) and (39) in (14), we get

$$\begin{aligned}
\mathbf{V}(\alpha) =& \int_0^\infty \left(e^{-\frac{1}{2}}\mathbf{u}_N\mathbf{v}_N^T + \sum_{i=1}^{N-1} e^{t(\alpha(-1-\lambda_i)-\frac{1}{2})}\mathbf{u}_i\mathbf{v}_i^T\right)\left(\sum_{k=1}^{N-1}\frac{1+\lambda_k}{1-\lambda_k}\mathbf{u}_k\mathbf{u}_k^T\right)\left(e^{-\frac{1}{2}}\mathbf{u}_N\mathbf{v}_N^T + \sum_{i=1}^{N-1} e^{t(\alpha(-1-\lambda_i)-\frac{1}{2})}\mathbf{u}_i\mathbf{v}_i^T\right)^T dt \\
=& \int_0^\infty \left(\sum_{i=1}^{N-1} e^{t(\alpha(-1-\lambda_i)-1/2)}\mathbf{u}_i\mathbf{v}_i^T\right)\left(\sum_{k=1}^{N-1}\frac{1+\lambda_k}{1-\lambda_k}\mathbf{u}_k\mathbf{u}_k^T\right)\left(\sum_{i=1}^{N-1} e^{t(\alpha(-1-\lambda_i)-1/2)}\mathbf{u}_i\mathbf{v}_i^T\right)^T dt \\
=& \int_0^\infty \sum_{i=1}^{N-1} e^{2t(\alpha(-1-\lambda_i)-1/2)}\frac{1+\lambda_i}{1-\lambda_i}\mathbf{u}_i\mathbf{u}_i^T dt = \sum_{i=1}^{N-1}\frac{1}{2\alpha(1+\lambda_i)+1}\cdot\frac{1+\lambda_i}{1-\lambda_i}\mathbf{u}_i\mathbf{u}_i^T
\end{aligned}$$

[6]The map $\mathbf{f}$ can be thought as a function to be estimated by sampled at X_{n+1} and corresponding to the update rule evaluated at $\mathbf{x}_n = \mathbf{d}$, that is, $\mathbf{F}_{X_{n+1},\cdot} = \mathbf{f}(X_{n+1}) = \boldsymbol{\delta}_{X_{n+1}} - \mathbf{d}$.

where the first three equalities follow from orthonormality of the left and right eigenvectors. The last equality comes from swapping the summation with the integral,[7] and evaluating the latter. This completes the proof. □

Proof of Corollary 4.3. For any $\alpha > 0$ and any vector $\mathbf{x} \in \mathbb{R}^N$, we have

$$\mathbf{x}^T\mathbf{V}(\alpha)\mathbf{x} = \sum_{i=1}^{N-1} \frac{1}{2\alpha(1+\lambda_i)+1} \cdot \frac{1+\lambda_i}{1-\lambda_i}\mathbf{x}^T\mathbf{u}_i\mathbf{u}_i^T\mathbf{x} < \frac{1+\lambda_i}{1-\lambda_i}\mathbf{x}^T\mathbf{u}_i\mathbf{u}_i^T\mathbf{x} = \mathbf{x}^T\mathbf{V}(0)\mathbf{x},$$

where the inequality is because $\lambda_i \in (-1, 1)$, and as a result, $\alpha(-\lambda_i - 1) < 0$ for all $i \in \{1, \cdots, N-1\}$. In fact, the ordering is monotone in $\alpha > 0$. This completes the proof. □

Proof of Corollary 4.6. Note that every L^p norm $\|\cdot\|_p$ is a continuous and bounded function on Σ. The implication then follows by a direct application of the continuous mapping theorem (see Van der Vaart, 2000, Theorem 2.3) to the CLT shown in Theorem (4.2). □

Proof of Corollary 4.7. Recall that in Section 2, we had redefined $\mathbf{x}_n$ as $\mathbf{x}_n \triangleq \frac{1}{n+1}(\mathbf{x}_0 + \sum_{k=1}^n \boldsymbol{\delta}_{X_k})$, where $\mathbf{x}_0 = \boldsymbol{\nu} \in \text{Int}(\Sigma)$ was set a-priori. We redefine the actual empirical distribution of the process as $\hat{\mathbf{x}}_n \triangleq \frac{1}{n}\sum_{k=1}^n \boldsymbol{\delta}_{X_k}$, and rewrite $\mathbf{x}_n$ as

$$\mathbf{x}_n = \frac{1}{n+1}\mathbf{x}_0 + \frac{n}{n+1}\hat{\mathbf{x}}_n \tag{40}$$

It is enough to show the almost sure convergence and CLT result for $\hat{\mathbf{x}}_n$, since the result for $\psi_n(g) = \mathbf{g}^T\hat{\mathbf{x}}_n$ follows from the boundedness assumption for $\mathbf{g}$ (which ensures square summability). From Theorem 4.1, and because $\mathbf{x}_0/n+1 \to \mathbf{0}$, as $n \to \infty$, we have that $\mathbf{x}_n n/(n+1) \to \boldsymbol{\mu}$ almost surely. Multiplying $\mathbf{x}_n n/(n+1)$ by $(n+1)/n$, and since $(n+1)/n \to 1 < \infty$, we obtain that $\mathbf{x}_n \to \boldsymbol{\mu}$ almost surely.

From the CLT in Theorem 4.2, we have $\sqrt{n+1}(\mathbf{x}_n - \boldsymbol{\mu}) \to \mathcal{N}(0, \mathbf{V}(\alpha))$ in distribution. We will again break down $\mathbf{x}_n$ to obtain

$$\frac{n}{\sqrt{n+1}}\hat{\mathbf{x}}_n - \sqrt{n+1}\boldsymbol{\mu} \xrightarrow[n\to\infty]{dist} \mathcal{N}(0, \mathbf{V}(\alpha)),$$

where the $\mathbf{x}_0$ term is not present since $\frac{1}{\sqrt{n+1}}\mathbf{x}_0 \to \mathbf{0}$. We multiply the above by $\sqrt{n+1}/\sqrt{n}$ to get

$$\frac{\sqrt{n+1}}{\sqrt{n}}\left(\frac{n}{\sqrt{n+1}}\hat{\mathbf{x}}_n - \sqrt{n+1}\boldsymbol{\mu}\right) \xrightarrow[n\to\infty]{dist} \mathcal{N}(0, \mathbf{V}(\alpha)),$$

where the convergence holds because $\sqrt{n+1}/\sqrt{n} \to 1 < \infty$ and by applying Slutsky's theorem. Additionally, observe that

$$\frac{\sqrt{n+1}}{\sqrt{n}}\left(\frac{n}{\sqrt{n+1}}\hat{\mathbf{x}}_n - \sqrt{n+1}\boldsymbol{\mu}\right) = \sqrt{n}(\mathbf{x} - \boldsymbol{\mu}) + \frac{1}{n}\boldsymbol{\mu},$$

where the term $\frac{1}{n}\boldsymbol{\mu} \to \mathbf{0}$, implying that $\sqrt{n}(\mathbf{x} - \boldsymbol{\mu}) \xrightarrow[n\to\infty]{dist} \mathcal{N}(0, \mathbf{V}(\alpha))$. This completes the proof. □

D. Scale invariance of SRRW transition probabilities with polynomial form of $r_{\mu_i}(x_i)$

In Section 1, we briefly mentioned that the polynomial form of $r_{\mu_i}(x_i)$ as introduced in (3) for all $i \in \mathcal{N}$ is the only possible choice where transition probabilites $\mathbf{K}_{ij}[\mathbf{x}]$ are *scale invariant* - the entries of the target distribution μ_i and empirical measure x_i only need to be known up to a constant multiple for neighboring nodes of the random walker's position at each time step. This is equivalent to saying that for any $\mu_i \in \mathbb{R}$ (for any $x_i \in \mathbb{R}$), we have $r_{\mu_i}(Cx_i) = g(C)r_{\mu_i}(x_i)$ (we have $r_{C\mu_i}(x_i) = g(C)r_{C\mu_i}(x_i)$) for some function $g : \mathbb{R} \to \mathbb{R}$. Indeed, observe that for any $i, j \in \mathcal{N}$ and $r_{\mu_i}(x_i)$ as discussed, we have

$$K_{ij}[C\mathbf{x}] = \frac{P_{ij}r_{\mu_j}(Cx_j)}{\sum_k P_{ik}r_{\mu_k}(Cx_k)} = \frac{P_{ij}g(C)r_{\mu_j}(x_j)}{\sum_k P_{ik}g(C)r_{\mu_k}(x_k)} = \frac{P_{ij}r_{\mu_j}(x_j)}{\sum_k P_{ik}r_{\mu_k}(x_k)} = K_{ij}[\mathbf{x}].$$

The following result shows such scale-invariance is only a property of polynomial choice of $r_{\mu_i}(x_i)$ for all $i \in \mathcal{N}$.

[7] From Fubini's theorem, the order of summation and integrals can be swapped if the summands are all positive terms - which stands true for our case.

Proposition D.1. *For any function $r : \mathbb{R} \to \mathbb{R}$, there exists a function $g : \mathbb{R} \to \mathbb{R}$ such that*

$$r(C_1 x) = g(C_1) r(x) \tag{41}$$

for any $C_1 \in \mathbb{R}$ if and only if $r(x)$ is of the form $r(x) = C_2 x^\alpha$ for any $C_2, \alpha \in \mathbb{R}$.

Proof. The reverse implication of the statement is true for any polynomial function $r : \mathbb{R} \to \mathbb{R}$ of type $r(x) = C_2 x^\alpha$ with $g(C_1) = C_1^\alpha$. We now prove the forward direction. Differentiating (41) on both sides, we get

$$C_1 r'(C_1 x) = g(C_1) r'(x) \tag{42}$$

and dividing the two sides of (41) by those of (42) yields

$$\frac{r(C_1 x)}{r'(C_1 x)} = C_1 \frac{r(x)}{r'(x)}. \tag{43}$$

Setting $f(x) \triangleq r(x)/r'(x)$ for all $x \in \mathbb{R}$, substituting in (43), and then differentiating once more gives us

$$f'(C_1 x) = f'(x) \tag{44}$$

for all $C_1, x \in \mathbb{R}$. This is only possible if $f : \mathbb{R} \to \mathbb{R}$ is a linear function, which by its definition is only possible if $r(x)$ is a polynomial function of the type $r(x) = C_2 x^\alpha$, for some $C_2, \alpha \in \mathbb{R}$. □

E. Discussion on Assumption (A1)

We first describe the modified stochastic approximation procedure with restarts of the process upon hitting the boundaries of a sequence of (expanding) compact subsets of $\text{Int}(\Sigma)$. Define a sequence of compact subsets $\{\mathcal{K}_n\}_{n \in \mathbb{N}_0}$ of $\text{Int}(\Sigma)$ such that $\mathcal{K}_n \subset \mathcal{K}_{n+1}$ for all $n \in \mathbb{N}_0$, and $\cup_{n \in \mathbb{N}_0} \mathcal{K}_n = \text{Int}(\Sigma)$. Let $\{\bar{\boldsymbol{\gamma}}^m\}_{m \in \mathbb{N}_0}$ be a family of step size sequences, where $\bar{\boldsymbol{\gamma}}^m \triangleq \{\gamma_{k,m}\}_{k \in \mathbb{N}_0}$ for all $m \in \mathbb{N}_0$, with $\gamma_{k,m} \triangleq \gamma_{k+m,0} \triangleq 1/(k+m+2)$, for all $k, m \in \mathbb{N}_0$.

Setting $\mathbf{x}_0 \in \mathcal{K}_0$, where $\mathcal{K}_0$ is the *active set*, and setting the step-size sequence to be $\bar{\boldsymbol{\gamma}}^0$, we run the iteration

$$\mathbf{x}_{n+1} = (1 - \gamma_{n,m}) \mathbf{x}_n + \gamma_{n,m} \boldsymbol{\delta}_{X_{n+1}}, \tag{45}$$

where $m = 0$ and $X_{n+1} \sim K_{X_n,\cdot}[\mathbf{x}_n]$, until the iterate leaves the active set $\mathcal{K}_0$. Upon this event (also called a truncation), we 'expand' the active set by incrementing its index and setting it to be $\mathcal{K}_1$, set the new step size sequence to be $\boldsymbol{\gamma}^1$, and restart the iteration (45) with these incremented active sets and step size sequences, and with a new initial point $\mathbf{x}_0 \in \mathcal{K}_0$, in an *i.i.d.* manner upon each restart. This process of truncation and restarts keeps repeating, and as part of the proof of our first order convergence results, we show that the number of restarts is always finite $\mathbb{P}_{\mathbf{x}_0, X_0}$ - almost surely.

This SA procedure with truncations can also be written more comprehensively. Let the step-size sequence be $\bar{\boldsymbol{\gamma}} \triangleq \{\gamma_k\}_{k \in \mathbb{N}_0}$, where $\gamma_k \triangleq \gamma_{k,0} = 1/(k+2)$. At each step $n \in \mathbb{N}_0$, let ς_n, κ_n and ν_n be *counters* keeping track of the step-size index, the index of the active set, and the number of iterations since the last truncation event, respectively. With $\mathbf{x}_0 \in \mathcal{K}_0$, $X_0 \in \mathcal{N}$ as before, and $\varsigma_0 = 0$, $\kappa_0 = 0$, and $\nu_0 = 0$, the SA procedure with truncations can be written as

$$\begin{aligned} &\text{set:} \quad \mathbf{x}_{n+\frac{1}{2}} = \mathbf{x}_n + \gamma_{\varsigma_n+1}(\boldsymbol{\delta}_{X_{n+1}} - \mathbf{x}_n), \\ &\text{update:} \quad (\mathbf{x}_{n+1}, \varsigma_{n+1}, \kappa_{n+1}, \nu_{n+1}) = \begin{cases} (\mathbf{x}_{n+\frac{1}{2}}, \varsigma_n+1, \kappa_n, \nu_n+1), & \text{if } \mathbf{x}_{n+\frac{1}{2}} \in \mathcal{K}_{\kappa_n} \\ (\mathbf{x}_0, \varsigma_n+1-\nu_n, \kappa_n+1, 0), & \text{otherwise.} \end{cases} \end{aligned} \tag{46}$$

Proposition E.1. *The iterates $\{\mathbf{x}_n\}_{n \geq 0}$ of (46) satisfy assumption (A1).*

Before stating the proof, we make the case for why our SRRW iterates in (9) are almost surely contained within compact subsets[8] of $\text{Int}(\Sigma)$, without the need of any truncation at boundaries of an increasing sequence of compact sets. Let the increasing sequence of compact sets be given by $\mathcal{K}_n \triangleq \{\mathbf{x} \in \text{Int}(\Sigma) | x_i \in \left[\frac{1}{n+M}, 1 - \frac{1}{n+M}\right]\}$, where M can be any positive real number. As explained in (Andrieu et al., 2005), the requirement that $\mathbf{x}_0 \in \mathcal{K}_0$ is under the condition that $\mathcal{K}_0$ is a subset

[8]The compact subset may depend on the sample path.

of a region where iterates eventually experience a *positive drift* towards the equilibrium point. However, the uniqueness of the $\boldsymbol{\mu} \in \text{Int}(\Sigma)$ as the fixed point of our mean-field ODE shown in Lemma 3.1, and the form of our *strict* Lyapunov function as in Lemma 3.2 allows us to get rid of this requirement, allowing $\mathcal{K}_0$ to be *any* compact subset of $\text{Int}(\Sigma)$. This allows us to choose the parameter M defining the size of $\mathcal{K}_0$ to be large enough so that $\mathcal{K}_0 \approx \text{Int}(\Sigma)$ and as a result, $\mathcal{K}_n \approx \text{Int}(\Sigma)$ for all $n \geq 0$. In this manner, the effect of increasing the truncation set is made nearly redundant, and are likely not the key factor to maintaining stability of iterates $\{\mathbf{x}_n\}_{n\geq 0}$ of (46).

Proof of Proposition E.1. We prove this result for any step size sequence $\{\gamma_n\}_{n\geq 0}$ satisfying

B1 $\sum_{k\geq 0} \gamma_k = \infty$, and $\sum_{k\geq 0} \gamma_k^{2-\epsilon} < \infty$ for some $\epsilon \in (0,1)$.

Which includes the step size $\gamma_n = \frac{1}{n+1}$ considered in our paper. The above assumption is only slightly stricter than the typical one where $\epsilon = 0$, such as (A4) in (Fort, 2015) and (D) in (Delyon, 2000) , and we show that ϵ need only be very small. Therefore in practice, B1 is nearly indistinguishable from (A4) in (Fort, 2015) and (D) in (Delyon, 2000). For our choice of step-size $\gamma_n = \frac{1}{n+1}$, there exists $\epsilon > 0$ small enough such that (γ_n, ϵ) satisfy B1.

We first introduce a sequence $\{\varepsilon_n\}$ where $\varepsilon_n = 2\gamma_n^\delta$ for some $\delta \in (0,1)$ (and thus $2\gamma_n < \varepsilon_n$). The condition for acceptance of $\mathbf{x}_{n+\frac{1}{2}}$ can then be rewritten as requiring $\|\mathbf{x}_{n+\frac{1}{2}} - \mathbf{x}_n\| < \epsilon_{\varsigma_n}$ along with $\mathbf{x}_{n+\frac{1}{2}} \in \mathcal{K}_{\kappa_n}$, where the former is trivially satisfied since $\|\mathbf{x}_{n+\frac{1}{2}} - \mathbf{x}_n\| \leq \gamma_{\varsigma_n}\|\delta_{X_{n+1}-\mathbf{x}_n}\| \leq 2\gamma_{\varsigma_n} < \varepsilon_{\varsigma_n}$. With this modification, update rule (46) is then a special case of the general SA algorithm described in Section 3.2 in (Andrieu et al., 2005). The rest of the proof will then be checking that the assumptions required for applying Theorem 5.4 in (Andrieu et al., 2005) are satisfied.

Assumption (A1) in (Andrieu et al., 2005) is satisfied with $V(\mathbf{x})$ in (11) as the choice of Lyapunov function. A1(i)–(iv) in (Andrieu et al., 2005) all follow from Lemma 3.2, coupled with the fact that $\mathbf{x}^* = \mathbf{d} \in \text{Int}(\Sigma)$ is the unique fixed point; the set of equilibria $\mathcal{L} = \{\mathbf{d}\}$ is a singleton and therefore a closed set with non-empty interior, and the constants M_0 and M_1 can be any real numbers such that $V(\mathbf{d}) < M_0 < M_1 < \infty$.

Assumption (A2) in (Andrieu et al., 2005) is naturally satisfied by the construction of our SA algorithm, since $K[\mathbf{x}]$ is irreducible for any $\mathbf{x} \in \text{Int}(\Sigma)$.

We now check the set of assumptions (DRI) in (Andrieu et al., 2005). The condition (DRI1) is satisfied by any ergodic Markov chain [cite], and therefore also by $K[\mathbf{x}]$ for any $\mathbf{x} \in \mathcal{K} \subset \text{Int}(\Sigma)$ with $V(i) = 1$ for all $i \in \mathcal{N}$, where $\mathcal{K}$ is any compact subset of $\text{Int}(\Sigma)$. Condition (DRI2) when translated to our setting requires checking for any compact $\mathcal{K}, \mathcal{K}' \subset \text{Int}(\Sigma)$ that

$$\sup_{\mathbf{x}\in\mathcal{K}} \|\boldsymbol{\delta}_i - \mathbf{x}\| \leq C_1, \quad \text{and} \quad \sup_{(\mathbf{x},\mathbf{y})\in\mathcal{K}\times\mathcal{K}'} \|\mathbf{y}-\mathbf{x}\|^{-\beta}\|\mathbf{y}-\mathbf{x}\| \leq C_1$$

for some $C \in \mathbb{R}$ and $\beta \in [0,1]$. This clearly holds true with $C_1 = 2$ and $\beta = 1$. The condition (DRI3) when translated to our setting requires showing for any $(\mathbf{x},\mathbf{y}) \in \mathcal{K}\times\mathcal{K}'$ that there exists $C_2 \in \mathbb{R}$ such that

$$\|\mathbf{K}[\mathbf{x}]\mathbf{u} - \mathbf{K}[\mathbf{y}]\mathbf{u}\| \leq C_2\|\mathbf{u}\|\|\mathbf{x}-\mathbf{y}\|^\beta, \quad \forall \mathbf{u} \in \mathbb{R}^N, \|\mathbf{u}\| < \infty.$$

This is again clearly holds with $\beta = 1$ and for some $C_2 < \infty$, since $\mathbf{K}[\cdot]$ is not just continuous but also Lipschitz in any compact subset of $\text{Int}(\Sigma)$. Thus, update rule (46) satisfies (DRI) which implies (A3) in (Andrieu et al., 2005).

In order to satisfy (A4) in (Andrieu et al., 2005), we need to show that the sequences $\{\gamma_n\}_{n\geq 0}$ and $\{\varepsilon_n\}_{n\geq 0}$ are non-increasing, positive, and satisfy $\sum_{k\geq 0}\gamma_k = \infty$, $\lim_{k\to\infty}\varepsilon_k = 0$, and

$$\sum_{k\geq 0}\left\{\gamma_k^2 + \gamma_k\epsilon^a + (\epsilon_k^{-1}\gamma_k)^p\right\} < \infty.$$

Here, we can set $0 < a < \beta = 1$ and $p \geq 2$ (these can be deduced from the drift conditions (DRI), as discussed in Section 6 of (Andrieu et al., 2005)). By setting $\varepsilon_n = 2\gamma_n^\delta$ for some small $\delta \in (0,1)$, the condition boils down to choosing $\alpha \in (0,1)$ such that

$$\sum_{k\geq 0}\gamma_k^{a\delta+1} < \infty.^{9}$$

[9]This implies summability of γ_k^2 term, while that of the $(\varepsilon_k^{-1}\gamma_k)^p$ is ensured by setting $p \geq 2$ to a large enough value.

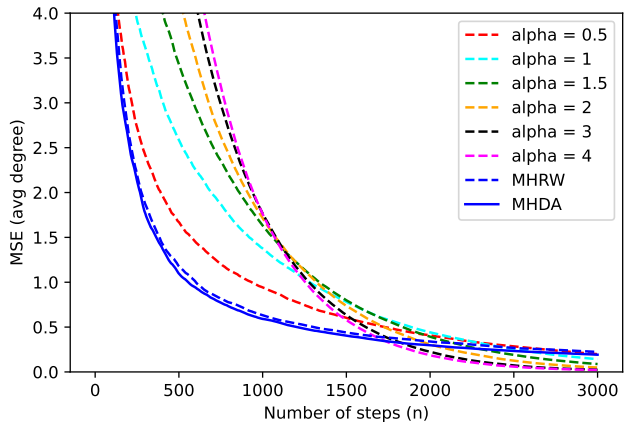
By setting $\epsilon = 1 - a\delta > 0$ in assumption B1, we can see that (A4) in (Andrieu et al., 2005) is satisfied. Note that $a\delta$ can be chosen to be very close to 1, which implies that in practice, B1 is no stricter than the square suitability assumption typically found in SA literature.

We can now apply Theorem 5.4 in (Andrieu et al., 2005) and the number of restarts are finite $\mathbb{P}_{\mathbf{x}_0, X_0}$-almost surely, implying that the iterate sequence will eventually remain in a compact subset of $\text{Int}(\Sigma)$. □

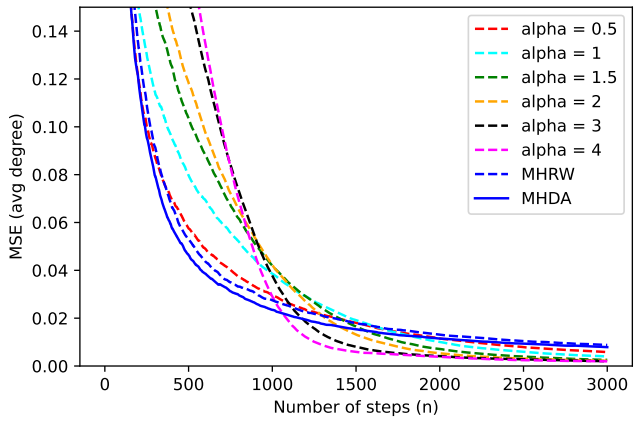
F. Additional numerical results

In this appendix, we provide supplementary numerical results to those in Section 5. We compare the sampling performance SRRW with the MHRW as its base chain, with the Metropolis-Hastings with delayed acceptance (MHDA) sampler introduced in (Lee et al., 2012). The MHDA works by reducing the inherent backtracking probability of MHRW by interacting with its most recently visited state (see Lee et al., 2012, Section 4.3 for the detailed definition of the transition probabilities). We provide our numerical results for two different graphs over the same set of nodes. As mentioned in the captions for Figures 2(a) and 2(b), the base MHRW chain for both these graphs have different mixing properties.

The simulation setup is similar to the one in Section 5, the numerical results are shown Figure 2 where we focus on the MSE of the estimators. For values of $\alpha > 0$ which are only moderately large, the SRRW significantly outperforms the MHDA sampler, showing the performance improvement in the asymptotic regime from interacting with the entire history of the random walker instead of only the most recently visited state. We also observe that performance benefits of SRRW with larger α kick in earlier when the the underlying base chain is faster mixing, which is the case on the Erdos-Renyi random graph in Figure 2(b).



(a) Convergence of $\psi_n(g)$ to the ground truth $\mathbf{g}^T\mathbf{1}/N$ for the wikiVote graph (889 nodes, 2914 edges, SLEM $= 0.99$). Base MHRW mixes slower.



(b) Convergence of $\psi_n(g)$ to the ground truth $\mathbf{g}^T\mathbf{1}/N$ for an Erdos-Renyi random graph (889 nodes, 3927 edges, SLEM $= 0.93$). Base MHRW mixes faster.

Figure 2. Simulations of the SRRW process for values of $\alpha \in [0, 4]$, where $\alpha = 0$ corresponds to *MHRW* - the underlying Metropolis-Hastings *base* chain, with no self-repellence properties.