

# Stochastic Stability and Time-Dependent Mutations\*

Maxwell Pak<sup>†</sup>  
Queen's University

November 2007

## Abstract

This paper considers stochastic stability analysis in evolutionary models with time-dependent mutations. It takes a class of time-homogeneous Markov models where the transition probabilities are approximately polynomial functions of the mutation parameter and allows the mutation parameter to decline to zero over time. The main result shows that as long as the mutation parameter converges to zero slowly enough and its variation is finite, the resulting time-inhomogeneous model has a limiting distribution regardless of the details of the mutation process. Moreover, a bound on the required rate of decline is explicitly expressed as a function of the minimum coradius of the limit sets and the transition probabilities within the minimum coradius set.

**Keywords:** Evolution, Learning, Stochastic Stability.

**JEL Classification:** C7, D8.

---

\*I am very grateful to Jim Bergin and Katsumi Shimotsu for helpful comments and discussions. I am also indebted to the associate editor and the referees for their valuable comments and suggestions.

<sup>†</sup>Department of Economics, Queen's University. Kingston, Ontario K7L 3N6, Canada. Phone: 613.533.2251. Fax: 613.533.6668. E-mail: pakm@econ.queensu.ca.

# 1 Introduction

Stochastic stability analysis has by now become a familiar tool in economics for studying dynamical systems that are modeled as a finite state space Markov chain. To get an overview of this technique, suppose a Markov transition matrix  $\bar{P}$  describing some dynamical system is specified. If  $\bar{P}$  is ergodic, the system would have a limiting distribution that uniquely describes its long-run behavior independently of any initial conditions.<sup>1</sup> Typically, however, the specified matrix  $\bar{P}$  is not ergodic; therefore, the long-run behavior of the system depends critically on its starting condition. For example, if the system starts with any of its invariant distributions as the initial distribution, then it would always be governed by that distribution. Moreover, if the system starts with some other arbitrary distribution, then the long-run behavior may not even converge.

In such cases, stochastic stability analysis is often used to select a particular invariant distribution as the “likely long-run distribution.” This usually involves adding a noise term, parameterized by a mutation parameter  $\varepsilon$ , to each transition probability to derive a perturbed transition matrix  $P(\varepsilon)$  that is ergodic. Since  $P(\varepsilon)$  is ergodic, it has a corresponding limiting distribution  $\pi(\varepsilon)$ . While  $\pi(\varepsilon)$  need not be an invariant distribution for the original system  $\bar{P}$ , the distribution  $\pi^*$ , which is obtained as the limit of  $\pi(\varepsilon)$  as  $\varepsilon$  goes to zero, is necessarily an invariant distribution for  $\bar{P}$ . In this regard,  $\pi^*$  can be interpreted as the invariant distribution of  $\bar{P}$  that is robust to small perturbations in the transition probabilities, and any state which is in the support of  $\pi^*$  is called a “stochastically stable state” or a “long-run equilibrium” in the literature. The distribution  $\pi^*$  will be called the stochastically stable distribution in this paper.

Stochastic stability has proven useful as a robustness criteria. However, care must be exercised in its use since the invariant distribution that is selected depends crucially on mutation rates. Bergin and Lipman [1] show that it is possible to select *any* invariant distribution as the stochastically stable distribution simply by allowing mutation rates to depend on the state of the system.

Given the sensitivity of the stochastic stability analysis to the specification of mutation rates across states, it is natural to ask how the analysis is affected when mutation rates are allowed to vary across time. Moreover, if mutations are interpreted as experimentation, then it may be descriptively unrealistic to assume that mutations are constant across time. Consequently, this assumption is abandoned and instead the following more natural model is considered here. As usual, assume that agents mostly follow some optimizing behavior specified by a transition matrix  $\bar{P}$ . However, assume now that while agents also experiment, they experiment less as they gain experience so that the mutation parameter goes to zero over time. Since transition probabilities now vary over time, the fact that  $P(\varepsilon)$  is ergodic for each  $\varepsilon$  does not guarantee that the resulting time-inhomogeneous system has a convergent

---

<sup>1</sup>A transition matrix  $P$  on countable set  $S$  is said to be ergodic if there exists a limiting distribution  $\pi$  on  $S$  such that  $\pi = \lim_{t \rightarrow \infty} \mu_0 (P)^t$  for all initial distribution  $\mu_0$  on  $S$ . A well-known result for Markov chains states that if  $P$  is aperiodic, irreducible, and non-null recurrent, then it is ergodic.

long-run behavior. Therefore, this paper studies the conditions on the mutation rates that guarantee that the resulting inhomogeneous system has a limiting distribution.

This question has been first studied in Robles [9], which considers Kandori, Mailath, and Rob [8] and Young [12] (KMR/Y) models where transition probabilities are polynomial functions of the mutation parameter. Robles allows the mutation parameter to decrease to zero monotonically and derives sufficient conditions for ergodicity. These conditions, which are based on the minimum coradius of the limit sets, have intuitive appeal. However, as the example in this paper shows, they do not guarantee ergodicity when the base dynamics  $\bar{P}$  has periodic cycles. More recently, Chen and Chow [2] consider  $2 \times 2$  games with both uniform and local matching rules and provide conditions for ergodicity that are similar to Robles's but without requiring monotonicity. In a related work, Sandholm and Pauzner [10] consider the effect of population growth in the KMR model. They show that if the mutation rate remains constant but the population grows at a sub-logarithmic rate, then the resulting dynamics is ergodic.<sup>2</sup> However, if the population growth is at least logarithmic and, in addition, the mutation parameter is taken to zero, then the long-run dynamics is completely determined by the initial conditions.

This paper extends the earlier results of Robles and Chen and Chow to any finite state space Markov models, including those with periodic cycles, whose transition probabilities are approximately polynomial functions of the mutation parameter. In particular, we take a class of time-homogeneous models that is slightly more restrictive than that given in Ellison [4] and allow the mutation parameter to decline to zero over time. The main result shows that if the mutation parameter declines slowly enough and its variation is finite, then the resulting time-inhomogeneous system has a limiting distribution that is independent of the initial conditions. This limiting distribution is necessarily  $\pi^*$ , the stochastically stable distribution of the underlying time-homogeneous system.

The remainder of the paper is organized as follows. The inhomogeneous model of evolution that incorporates time-dependent mutation is defined in Section 2. Section 3 presents the results. The main result of the paper, Theorem 3.9, provides the rate of decline that guarantees the existence of a limiting distribution. This rate is closely related to the minimum coradius of the limit sets and the transition probabilities within the limit set with the minimum coradius (hereafter, the minimum coradius set). The paper concludes in section 4.

## 2 The Model

The class of underlying evolutionary models that is considered here is given in Definition 2.1 below, which is adapted from Ellison [4]. As in Ellison's formulation, the definition restricts attention to transition probabilities that are approximately

---

<sup>2</sup>As Sandholm and Pauzner show, keeping the mutation parameter constant and increasing the population size in the KMR model is equivalent to decreasing the mutation parameter and keeping the population constant.

polynomial functions of the mutation parameter.<sup>3</sup>

Below and elsewhere, notation  $P_{zz'}$  denotes the transition probability from  $z$  to  $z'$  under  $P$ , and  $\pi_z$  denotes the probability  $\pi$  places on  $z$ . Also, the following asymptotic notations are used throughout. The ‘little-oh’ notation  $f(x) = o(g(x))$  as  $x \rightarrow c$  means  $f(x)/g(x) \rightarrow 0$  as  $x \rightarrow c$ . The ‘big-oh’ notation  $f(x) = O(g(x))$  as  $x \rightarrow c$  means there exists a constant  $k > 0$  such that  $|f(x)| \leq k|g(x)|$  for all  $x$  sufficiently close to  $c$ . Similarly,  $f(x) = \Omega(g(x))$  as  $x \rightarrow c$  means there exists a constant  $k > 0$  such that  $|f(x)| \geq k|g(x)|$  for all  $x$  sufficiently close to  $c$ . Lastly,  $f(x) = \Theta(g(x))$  denotes  $f(x) = O(g(x))$  and  $f(x) = \Omega(g(x))$ .

**Definition 2.1.** A *homogeneous model of evolution* is a transition matrix  $\bar{P}$  and a family of transition matrices  $\{P(\varepsilon) : \varepsilon \in (0, \bar{\varepsilon}]\}$  on a finite set  $S$  such that

1.  $P(\varepsilon)$  is ergodic for all  $\varepsilon \in (0, \bar{\varepsilon}]$ ,
2.  $\lim_{\varepsilon \rightarrow 0} P(\varepsilon) = \bar{P}$ , and
3. there exists a cost function  $c : S \times S \rightarrow \mathbb{Z}_+ \cup \{+\infty\}$  and a family of Lipschitz functions  $\{f_{zz'}\}$  on  $[0, \bar{\varepsilon}]$  such that

$$P_{zz'}(\varepsilon) = k_{zz'} \varepsilon^{c(z, z')} (1 + f_{zz'}(\varepsilon)),$$

where  $k_{zz'} > 0$  and  $f_{zz'}(\varepsilon)$  is  $o(1)$  as  $\varepsilon \rightarrow 0$ .

**Remark 2.2.** This definition differs from Ellison [4] in two regards. First, Ellison allows the leading exponent  $c(z, z')$  to be extended non-negative reals,  $c : S \times S \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ , rather than just extended non-negative integers. Second, Ellison requires  $f_{zz'}(\varepsilon)$  to be merely continuous and not Lipschitz continuous. The additional restrictions are needed only because mutation rates are not assumed to decrease monotonically. As seen in the proof of Theorem 3.9, these assumptions place limits on how much the invariant distributions for the underlying homogeneous model fluctuate when the mutation parameter changes.

In a typical application, the transition matrix  $\bar{P}$  governs the evolution of some base dynamical system while, for each  $\varepsilon$ ,  $P(\varepsilon)$  governs the evolution of the system when the transition probabilities are perturbed according to mutation parameter  $\varepsilon$ . Let  $P^{(t)}(\varepsilon)$  denote the  $t$ -step transition matrix for  $P(\varepsilon)$ . For any initial distribution  $\mu_0$  on  $S$ , the distribution of the states at time  $t$  is given by  $\mu_0 P^{(t)}(\varepsilon)$ . Since  $P(\varepsilon)$  is assumed to be ergodic, there exists a limiting distribution  $\pi(\varepsilon)$  such that for every  $\mu_0$ ,

$$\lim_{t \rightarrow \infty} \mu_0 P^{(t)}(\varepsilon) = \lim_{t \rightarrow \infty} \mu_0 (P(\varepsilon))^t = \pi(\varepsilon).$$

While  $\pi(\varepsilon)$  need not be the limiting distribution or even an invariant distribution for the base system  $\bar{P}$ , the following well-known result states that  $\pi(\varepsilon)$  does converge to an invariant distribution of  $\bar{P}$  as  $\varepsilon \rightarrow 0$ .

---

<sup>3</sup>Since any continuous function on a compact interval can be uniformly approximated by a polynomial, these include most transition probabilities of practical interest.

**Theorem 2.3.**<sup>4</sup> For any homogeneous model of evolution, the limit

$$\pi^* \equiv \lim_{\varepsilon \rightarrow 0} \pi(\varepsilon) = \lim_{\varepsilon \rightarrow 0} \left( \lim_{t \rightarrow \infty} \mu_0 P^{(t)}(\varepsilon) \right) \quad (1)$$

exists and does not depend on  $\mu_0$ . Moreover, the limit  $\pi^*$  is an invariant distribution for  $\bar{P}$ .

If  $\bar{P}$  is ergodic, then the base dynamics already has a limiting distribution and, hence, a unique invariant distribution. By the above theorem, it must be  $\pi^*$ . So, adding mutation to the base model and letting it go to zero adds nothing to the analysis. Therefore,  $\bar{P}$  is assumed to be non-ergodic for the remainder of the paper.

Consider now a model of evolution in which the mutation parameter  $\varepsilon$  is allowed to vary over time. In particular, let  $\hat{\varepsilon} : \mathbb{Z}_+ \rightarrow (0, \bar{\varepsilon}]$  be a sequence converging to zero and let the transition from time  $t$  to time  $t+1$  be governed by  $P(\hat{\varepsilon}(t))$ . Then the sequence of transition matrices  $\{P(\hat{\varepsilon}(t)) : t \in \mathbb{Z}_+\}$  defines a single time-inhomogeneous Markov chain on  $S$ . For all  $t \in \mathbb{Z}_+$  and  $t' > t$ , let  $P^{(t,t')} \equiv \prod_{k=t}^{t'-1} P(\hat{\varepsilon}(k))$  denote the matrix of transition probabilities from time  $t$  to time  $t'$ .

**Definition 2.4.** An *inhomogeneous model of evolution* derived from a homogeneous model of evolution  $(\bar{P}, \{P(\varepsilon) : \varepsilon \in (0, \bar{\varepsilon}]\})$  is a sequence of transition matrices  $\{P(\hat{\varepsilon}(t)) : t \in \mathbb{Z}_+\} \subset \{P(\varepsilon) : \varepsilon \in (0, \bar{\varepsilon}]\}$  such that  $\hat{\varepsilon}(t) \rightarrow 0$ .

For any initial distribution  $\mu_0$ , the distribution of the states at time  $t'$  under the inhomogeneous dynamics is given by  $\mu_0 P^{(0,t')}$  as in the homogeneous model. Unlike the homogeneous model, however, the fact that  $P(\hat{\varepsilon}(t))$  is ergodic for all  $t$  does not guarantee that  $\mu_0 P^{(0,t')}$  converges to a limiting distribution as  $t' \rightarrow \infty$ . In fact, without further restrictions on  $\hat{\varepsilon}(t)$ ,

$$\lim_{t' \rightarrow \infty} \mu_0 P^{(0,t')} = \lim_{t' \rightarrow \infty} \mu_0 \prod_{t=0}^{t'-1} P(\hat{\varepsilon}(t)) = \mu_0 \left( \lim_{t' \rightarrow \infty} \prod_{t=0}^{t'-1} P(\hat{\varepsilon}(t)) \right)$$

may not even exist, let alone be the same for all  $\mu_0$ .

If the above limit is to exist and be the same for every  $\mu_0$ , then  $\prod_{t=0}^{t'-1} P(\hat{\varepsilon}(t))$  must converge to a transition matrix whose rows are all identical. Our main result, Theorem 3.9, shows that this occurs if (i)  $\hat{\varepsilon}(t)$  goes to zero at an asymptotic rate of  $t^{-\frac{1}{\gamma}}$  or slower, where  $\gamma$  is a constant that depends on the details of the underlying homogeneous model, and (ii) the variation of  $\hat{\varepsilon}(t)$  is finite.

Intuitively, the rate at which  $\hat{\varepsilon}(t)$  approaches zero matters because the one-step transition matrix at time  $t$ ,  $P(\hat{\varepsilon}(t))$ , converges to a non-ergodic transition matrix  $\bar{P}$  as  $\hat{\varepsilon}(t) \rightarrow 0$ . If  $P(\hat{\varepsilon}(t))$  approaches  $\bar{P}$  too quickly, or equivalently if  $\hat{\varepsilon}$  approaches zero too quickly, then the inhomogeneous dynamics will not have the chance to settle down to a limiting behavior. In contrast, suppose  $\varepsilon_1 \in (0, \bar{\varepsilon}]$  is fixed and the system is allowed to evolve according to  $P(\varepsilon_1)$  for  $t_1$  periods. Since  $P(\varepsilon_1)$  has a limiting

---

<sup>4</sup>See, for example, Kandori, Mailath, and Rob [8], Young [12], Bergin and Lipman [1], or Ellison [4].

distribution  $\pi(\varepsilon_1)$ , the behavior of the system will be close to  $\pi(\varepsilon_1)$  at time  $t_1$  if  $t_1$  is sufficiently large. Now, suppose at time  $t_1$  the mutation parameter drops to  $\varepsilon_2 < \varepsilon_1$  and the system evolves for another  $t_2$  periods. Again, if  $t_2$  is sufficiently large, then the system's behavior will be close to  $\pi(\varepsilon_2)$  at time  $t_1 + t_2$ . It is not difficult to imagine that if the mutation parameter continues to drop in this manner, then the behavior of the system will eventually converge to  $\pi^*$ .

In fact, decreasing the mutation parameter in this manner provides a straightforward method for deriving a time-inhomogeneous model of evolution that has a convergent limiting behavior. However, the artificial nature of the stepwise decline would make the model unsatisfactory. Our result shows that the mutation parameter decreasing to zero in this stepwise or even monotonic fashion is not necessary.

### 3 Results

For a Markov chain to have a limiting distribution that is independent of its initial distribution, it must necessarily have a “loss of memory” property. That is, as the chain runs, the probability of the chain visiting a given state should become increasingly less dependent on where the chain started from, and, in the limit, the chain should forget its initial state. The following definition captures this idea precisely.<sup>5</sup>

**Definition 3.1.** An inhomogeneous Markov chain on finite state space  $S$  is *weakly ergodic* if for all  $t \in \mathbb{Z}_+$  and  $z, z', z'' \in S$ ,

$$\lim_{t' \rightarrow \infty} \left| P_{zz''}^{(t,t')} - P_{z'z''}^{(t,t')} \right| = 0.$$

Although this definition requires the chain to eventually forget its initial condition, it does not require the chain's long-run behavior to actually converge. When a chain has this convergence property in addition to the loss of memory property, it is said to be strongly ergodic.

**Definition 3.2.** An inhomogeneous Markov chain on finite state space  $S$  is *strongly ergodic* if there exists a probability distribution  $\pi$  on  $S$  such that for all  $t \in \mathbb{Z}_+$  and  $z, z' \in S$ ,

$$\lim_{t' \rightarrow \infty} P_{zz'}^{(t,t')} = \pi_{z'},$$

or, equivalently, for all  $t \in \mathbb{Z}_+$  and for any initial distribution  $\mu_0$  on  $S$ ,

$$\lim_{t' \rightarrow \infty} \mu_0 P^{(t,t')} = \pi.$$

While we are ultimately interested in the strong ergodicity of an inhomogeneous model of evolution, we explore the weak ergodicity first since it is a necessary condition for strong ergodicity. Doeblin's theorem below gives the necessary and sufficient condition for an inhomogeneous chain to be weakly ergodic. Define a scalar function

---

<sup>5</sup>The two definitions of ergodicity are adopted from Seneta [11].

$\tau_1$ , called an ergodic coefficient, from the space of transition matrices on  $S$  to  $[0, 1]$  by

$$\tau_1(Q) = 1 - \min_{z, z' \in S} \sum_{z'' \in S} \min\{Q_{zz''}, Q_{z'z''}\}.$$
<sup>6</sup>

**Theorem 3.3** (Doebelin).<sup>7</sup> *An inhomogeneous Markov chain is weakly ergodic if and only if there is a strictly increasing sequence of positive integers  $\{t_k : k = 0, 1, 2, \dots\}$ , with  $t_0 = 0$ , such that*

$$\sum_{k=0}^{\infty} \left(1 - \tau_1\left(P^{(t_k, t_{k+1})}\right)\right) = \infty.$$

To get an intuition for this result, note that if  $\tau_1(Q) = 0$ , then all the rows of  $Q$  must be identical. Thus, transition probabilities do not depend at all on the initial state in this extreme case. Therefore, it is not too difficult to see that a chain can lose its memory if and only if its long-run ergodic coefficient goes to zero; that is, for all  $t$ ,  $\tau_1(P^{(t, t')}) \rightarrow 0$  as  $t' \rightarrow \infty$ . Theorem 3.3 states that this happens as long as the chain's medium-run ergodic coefficient does not approach one too quickly.

To apply Doebelin's theorem, a method for bounding  $\tau_1$  is needed. Fix any homogeneous model of evolution and consider the Markov chain defined by the unperturbed transition matrix  $\bar{P}$ . By the well-known decomposition theorem, the state space  $S$  can be partitioned uniquely as

$$S = T \cup C_1 \cup C_2 \cup \dots \cup C_M,$$

where  $T$  is transient under  $\bar{P}$  and for all  $m = 1, \dots, M$ ,  $C_m$  is irreducible, closed, and non-null recurrent under  $\bar{P}$ .<sup>8</sup> Since  $T$  is transient and finite, the unperturbed chain must almost surely visit one of the  $C_m$ 's in finite time. Once that happens, the chain will stay forever in  $C_m$ , visiting every element of  $C_m$  infinitely often with probability one. For this reason, each set  $C_m$  is commonly referred to as a minimal absorbing set, or a limit set, and  $C = \bigcup_{m=1}^M C_m$  is the set of all the limit points of the unperturbed dynamics.<sup>9</sup> If  $M = 1$  and  $C_1$  is aperiodic, then  $\bar{P}$  has a limiting distribution.<sup>10</sup> So, in this case, adding mutation to  $\bar{P}$  and letting it go to zero adds nothing to the analysis. Therefore, it is assumed that  $M > 1$  or  $C_1$  is periodic for the remainder of the paper.

A path  $h$  from  $z$  to  $z'$  is a collection of ordered pairs  $\{(z_1, z_2), (z_2, z_3), \dots, (z_{n_h-1}, z_{n_h})\}$  such that  $z_1 = z$  and  $z_{n_h} = z'$ . Let  $\tilde{c}(h) = \sum_{n=1}^{n_h-1} c(z_n, z_{n+1})$  denote the cost of transitioning along path  $h$ , and let  $\tilde{d}(h) = n_h - 1$  denote the length of path  $h$ . For

<sup>6</sup>Formally,  $\tau_1$  is only one example of a proper coefficient of ergodicity; however, an ergodic coefficient will always mean  $\tau_1$  in this paper. For other examples, see Seneta [11].

<sup>7</sup>See, for example, Seneta [11] Theorem 4.8 or Isaacson and Madsen [7] Theorem V3.2.

<sup>8</sup>See, for example, Grimmett and Stirzaker [6]. The fact that  $1 \leq M < \infty$  and  $C_m$ 's are non-null recurrent follows from the finiteness of  $S$ .

<sup>9</sup>A state  $z$  is called a limit point of  $P$  if there exists an initial distribution  $\mu_0$  on  $S$  such that if the chain starts with  $\mu_0$  and runs according to  $P$ , then the chain visits  $z$  infinitely often almost surely.

<sup>10</sup>To see this, note that since  $T$  is transient and  $C_1$  is closed, the chain governed by  $\bar{P}$  will almost surely enter  $C_1$  in finite time and never leave. Since  $C_1$  is irreducible, non-null recurrent, and aperiodic, the chain will thereafter behave like an ergodic chain with the state space limited to  $C_1$ .

any  $z, z' \in S$ , the resistance from  $z$  to  $z'$ ,  $r(z, z')$ , is defined as the minimum cost of transitioning from  $z$  to  $z'$ . That is,  $r(z, z') = \min \{\tilde{c}(h) : h \text{ is a path from } z \text{ to } z'\}$ .

For any  $z' \in S$ , the coradius of a point set  $\{z'\}$  is defined as  $CR(\{z'\}) = \max_{z \neq z'} r(z, z')$ , and represents the maximum resistance incurred in transitioning to  $z'$  from another state. The definition is extended to an arbitrary  $A \subset S$  by

$$CR(A) = \max_{z \notin A} \min_{z' \in A} r(z, z').$$

The coradius essentially measures how difficult it is, or how much resistance must be overcome, to transition into a set. In particular, if  $CR(A) < CR(S \setminus A)$ , then it is easier to transition into  $A$  than out of  $A$ . To keep the definition of coradius consistent with its intended meaning, set  $CR(S) = 0$  and  $CR(\emptyset) = \infty$ .

### 3.1 Coradius and Weak Ergodicity

The connection between the coradii of the homogeneous model of evolution and the ergodicity of the inhomogeneous model is most direct when the minimum coradius set is a singleton. To see this, suppose the minimum coradius set is  $\{\bar{z}\}$  and  $\beta$  is its coradius. For each  $z \neq \bar{z}$ , let  $d(z, \bar{z})$  be the length of a minimum cost path from  $z$  to  $\bar{z}$ . Let  $\hat{z}$  be the state that achieves the minimum coradius so that  $r(\hat{z}, \bar{z}) = \beta \geq r(z, \bar{z})$  for all  $z \in S$ . The choice of  $\hat{z}$  implies that the probability of transitioning from  $z$  to  $\bar{z}$  in exactly  $d(z, \bar{z})$  steps is asymptotically greater than or equal to the probability of transitioning from  $\hat{z}$  to  $\bar{z}$  in  $d(\hat{z}, \bar{z})$  steps. Moreover, once the chain reaches  $\bar{z}$ , it can continue to stay at  $\bar{z}$  with zero cost because  $c(\bar{z}, \bar{z})$  must be zero if  $\{\bar{z}\}$  is a limit set. So, for any  $n$  large enough,

$$P_{z\bar{z}}^{(n)}(\varepsilon) = \Theta \left( P_{z\bar{z}}^{(d(z, \bar{z}))}(\varepsilon) \right) = \Omega \left( P_{\hat{z}\bar{z}}^{(d(\hat{z}, \bar{z}))}(\varepsilon) \right) = \Omega \left( \varepsilon^\beta \right)$$

for every  $z \in S$ . Therefore,  $\varepsilon^\beta$  provides a lower bound on  $P_{z\bar{z}}^{(n)}(\varepsilon)$  that is independent of  $z$ .

This estimate can in turn be used to bound the ergodic coefficient. By taking  $t_{k+1} - t_k$  large enough, we have

$$\sum_{z'' \in S} \min \left\{ P_{zz''}^{(t_k, t_{k+1})}, P_{z'z''}^{(t_k, t_{k+1})} \right\} \geq \min \left\{ P_{z\bar{z}}^{(t_k, t_{k+1})}, P_{z'\bar{z}}^{(t_k, t_{k+1})} \right\} = \Omega \left( \hat{\varepsilon}(t_k)^\beta \right).$$

Therefore,

$$1 - \tau_1 \left( P^{(t_k, t_{k+1})} \right) = \Omega \left( \hat{\varepsilon}(t_k)^\beta \right).$$

Doebelin's theorem can now be readily applied to obtain a sequence  $\{\hat{\varepsilon}(t)\}$  that yields weak ergodicity.

This observation lies at the heart of Robles's earlier result which states that in KMR/Y models any monotonically decreasing sequence of mutation rates  $\hat{\varepsilon}(t)$  satisfying  $\sum_{t=0}^{\infty} \hat{\varepsilon}(t)^\beta = \infty$  yields a strongly ergodic inhomogeneous model.<sup>11</sup> However,

<sup>11</sup>See, Robles [9] Propositions 4.1 and 5.2.



this direct relationship between the coradius of the limit sets and the ergodicity breaks down when the minimum coradius set has periodic cycles. A key argument used above in deriving a bound on  $P_{z\bar{z}}^{(n)}(\varepsilon)$  that is independent of  $z$  had been the fact that once  $\bar{z}$  is reached, the chain can continue to stay there at zero cost. This need not hold if the limit set is periodic. In such cases, whether  $P_{z\bar{z}}^{(n)}(\varepsilon)$  is asymptotically greater than or equal to  $\varepsilon^\beta$  may depend on  $z$  no matter how large  $n$  is. Therefore, the bound identified by Robles may not be sufficient, as illustrated in the following example.

**Example 3.4.** Consider a model in which two players repeatedly play the following stage game.

		player 2	
		A	B
player 1	A	2, 1	0, 0
	B	0, 0	2, 1

Assume that the players are myopic so that in each period they simply best respond to the opponent's strategy in the previous period. Restricting attention to pure strategies, the evolution of the play can be described by a Markov chain on the state space  $S = \{AA, AB, BA, BB\}$ , where, for example,  $AB$  denotes player 1 playing  $A$  and player 2 playing  $B$ . The transition matrix is

$$\bar{P} = \left( \begin{array}{c|cccc} & AA & AB & BA & BB \\ \hline AA & 1 & 0 & 0 & 0 \\ AB & 0 & 0 & 1 & 0 \\ BA & 0 & 1 & 0 & 0 \\ BB & 0 & 0 & 0 & 1 \end{array} \right).$$

Since this chain is not ergodic, it has no limiting distribution.

Now, suppose that the players mostly best respond but sometimes they also experiment by not playing the best response. Suppose further that each player experiments more when she receives lower payoff than her opponent.<sup>12</sup> In particular, letting  $s_t^i$  denote player  $i$ 's strategy in period  $t$  and  $\pi_t^i$  denote player  $i$ 's payoff, assume that

$$s_{t+1}^i = \left\{ \begin{array}{l} \bullet \text{ If } \pi_t^i \geq \pi_t^j: \\ \quad \text{with probability } 1 - \varepsilon^2, \text{ play the best response to } s_t^j, \\ \quad \text{with probability } \varepsilon^2, \text{ play the strategy that is not the best response.} \\ \bullet \text{ If } \pi_t^i < \pi_t^j: \\ \quad \text{with probability } 1 - \varepsilon, \text{ play the best response to } s_t^j, \\ \quad \text{with probability } \varepsilon, \text{ play the strategy that is not the best response.} \end{array} \right.$$

---

<sup>12</sup>While this example uses state-dependent mutations, their use is not necessary. An example using state-independent mutations can be readily constructed, albeit at the cost of requiring a larger transition matrix.

The transition matrix corresponding to this perturbed dynamics has the following asymptotic behavior.

$$P(\varepsilon) = \Theta \left( \left( \begin{array}{c|cccc} & AA & AB & BA & BB \\ \hline AA & 1 & \varepsilon & \varepsilon^2 & \varepsilon^3 \\ AB & \varepsilon^2 & \varepsilon^4 & 1 & \varepsilon^2 \\ BA & \varepsilon^2 & 1 & \varepsilon^4 & \varepsilon^2 \\ BB & \varepsilon^3 & \varepsilon^2 & \varepsilon & 1 \end{array} \right) \right) \text{ as } \varepsilon \rightarrow 0.$$

The limit sets of  $\bar{P}$  are  $\{AA\}$ ,  $\{AB, BA\}$ , and  $\{BB\}$ . Their coradii under  $P(\varepsilon)$  are

$$CR(\{AA\}) = CR(\{BB\}) = 3 > 1 = CR(\{AB, BA\}),$$

and the minimum coradius,  $\beta$ , is 1.

Let  $\hat{\varepsilon}(t) = \frac{1}{t+2}$  for all  $t \in \mathbb{Z}_+$ . Then  $\hat{\varepsilon}(t) \downarrow 0$  and

$$\sum_{t=0}^{\infty} \hat{\varepsilon}(t)^\beta = \sum_{t=0}^{\infty} \frac{1}{t+2} = \infty.$$

Therefore,  $\hat{\varepsilon}(t)$  satisfies the conditions of Robles's [9] Propositions 4.1 and 5.2. However, as we show below, the corresponding inhomogeneous chain is not weakly ergodic.

Let  $n > 1$  be odd. A least cost path from  $AA$  to  $z \in \{AA, AB, BB\}$  in  $n$ -steps is the path

$$((AA, AA), (AA, AA), \dots, (AA, AA), (AA, z)),$$

and from  $AA$  to  $BA$  is

$$((AA, AA), (AA, AA), \dots, (AA, AA), (AA, AB), (AB, BA)).$$

A least cost path from  $AB$  to  $z \in \{AA, BB\}$  is the path

$$((AB, z), (z, z), \dots, (z, z), (z, z)),$$

from  $AB$  to  $AB$  is

$$((AB, AA), (AA, AA), \dots, (AA, AA), (AA, AB)),$$

and from  $AB$  to  $BA$  is

$$((AB, BA), (BA, AB), (AB, BA), \dots, (BA, AB), (AB, BA)).$$

Similar analysis for the least cost paths from  $BA$  to  $z \in S$  and  $BB$  to  $z \in S$  yields the following asymptotics for the  $n$ -step transition matrix when  $n > 1$  is odd.

$$P^{(n)}(\varepsilon) = \Theta \left( \left( \begin{array}{cccc} 1 & \varepsilon & \varepsilon & \varepsilon^3 \\ \varepsilon^2 & \varepsilon^3 & 1 & \varepsilon^2 \\ \varepsilon^2 & 1 & \varepsilon^3 & \varepsilon^2 \\ \varepsilon^3 & \varepsilon & \varepsilon & 1 \end{array} \right) \right) \text{ as } \varepsilon \rightarrow 0.$$

When  $n$  is even, similar analysis yields the following.

$$P^{(n)}(\varepsilon) = \Theta \left( \left( \begin{array}{cccc} 1 & \varepsilon & \varepsilon & \varepsilon^3 \\ \varepsilon^2 & 1 & \varepsilon^3 & \varepsilon^2 \\ \varepsilon^2 & \varepsilon^3 & 1 & \varepsilon^2 \\ \varepsilon^3 & \varepsilon & \varepsilon & 1 \end{array} \right) \right) \text{ as } \varepsilon \rightarrow 0.$$

Finally, when  $n = 1$ ,

$$P^{(n)}(\varepsilon) = P(\varepsilon) = \Theta \left( \left( \begin{array}{cccc} 1 & \varepsilon & \varepsilon^2 & \varepsilon^3 \\ \varepsilon^2 & \varepsilon^4 & 1 & \varepsilon^2 \\ \varepsilon^2 & 1 & \varepsilon^4 & \varepsilon^2 \\ \varepsilon^3 & \varepsilon^2 & \varepsilon & 1 \end{array} \right) \right) \text{ as } \varepsilon \rightarrow 0.$$

So, for any strictly increasing subsequence  $\{t_k : k \in \mathbb{Z}_+\}$ ,

$$\begin{aligned} \min_{z, z' \in S} \sum_{z'' \in S} \min \left\{ P_{zz''}^{(t_k, t_{k+1})}, P_{z'z''}^{(t_k, t_{k+1})} \right\} &= \Theta \left( \sum_{z'' \in S} \min \left\{ P_{AB, z''}^{(t_k, t_{k+1})}, P_{BA, z''}^{(t_k, t_{k+1})} \right\} \right) \\ &= \Theta \left( \hat{\varepsilon}(t_k)^2 \right) \\ &= \Theta \left( \frac{1}{(t_k)^2} \right) \\ &= O \left( \frac{1}{k^2} \right) \text{ as } k \rightarrow \infty. \end{aligned}$$

Therefore,

$$\sum_{k=0}^{\infty} \left( 1 - \tau_1 \left( P^{(t_k, t_{k+1})} \right) \right) < \infty,$$

and the chain is not weakly ergodic by Doeblin's theorem.<sup>13</sup>

□

### 3.2 Weak Ergodicity

As seen in the above example, the essence of the problem in the periodic case is as follows. Suppose a periodic limit set  $C_m$  is reached at time  $t$ . Since  $C_m$  is a limit set of  $\bar{P}$ , the chain can continue to stay in  $C_m$  at zero cost. However, because  $C_m$  is periodic under  $\bar{P}$ , whether the chain can visit a particular  $z' \in C_m$  in  $n$  steps without incurring further cost depends on the chain's location at time  $t$ , no matter

<sup>13</sup>Although the argument given in this example uses the formalism of ergodic coefficient, the insufficiency of Robles's conditions here can also be understood in terms of Borel-Cantelli lemma. Given a sequence of events  $\{A_t : t = 0, 1, 2, \dots\}$ , Borel-Cantelli lemma states that the probability of events  $A_t$ 's occurring infinitely often is zero if  $\sum_{t=0}^{\infty} \text{Prob}(A_t) < \infty$ . For a chain to have a limiting distribution that is independent of its initial condition, it must not have any cycles that occur with probability one. In the example, this necessarily means that the chain must transition from state  $AB$  to non- $BA$  states infinitely often. At time  $t$ , the probability of this event is of order  $\hat{\varepsilon}(t)^2 = \frac{1}{(t+2)^2}$ . Since  $\sum_{t=0}^{\infty} \frac{1}{(t+2)^2} < \infty$ , the probability of transitioning from  $AB$  to non- $BA$  states infinitely often is zero by Borel-Cantelli lemma. Therefore, the chain cannot be weakly ergodic.

how large  $n$  is. Therefore, deriving a bound on  $P_{z\bar{z}}^{(n)}(\varepsilon)$  that is independent of  $z$  requires introducing an adjustment factor that accounts for the cost of transitioning within the limit set.

Abusing notation slightly, let  $r(z, z', n)$  be the minimum cost of transitioning from  $z$  to  $z'$  in exactly  $n$  steps under  $P(\varepsilon)$ . For each  $z' \in C_m$ , let

$$\alpha(\{z'\}) = \min_{n \geq 1} \max_{z \in C_m} r(z, z', n).^{14}$$

Extend the definition to the entire  $C_m$  by letting

$$\alpha(C_m) = \min_{z' \in C_m} \alpha(\{z'\}).$$

Next, let

$$\gamma = \min_{z' \in C} CR(\{z'\}) + \alpha(\{z'\}).$$

For a homogeneous model of evolution,  $r(z, z') = 0$  for all  $z, z' \in C_m$  so that  $CR(\{z'\}) = CR(C_m)$  for all  $z' \in C_m$ .<sup>15</sup> Therefore,  $\gamma$  can also be found by minimizing over the limit sets rather than individual limit points, which may be more convenient in some applications.

$$\begin{aligned} \gamma &= \min_{z' \in C} CR(\{z'\}) + \alpha(\{z'\}) \\ &= \min_{C_m \in \{C_1, \dots, C_M\}} \left( CR(C_m) + \min_{z' \in C_m} \alpha(\{z'\}) \right) \\ &= \min_{C_m \in \{C_1, \dots, C_M\}} CR(C_m) + \alpha(C_m). \end{aligned}$$

The following lemma shows that if the mutation parameter decreases at a rate equal to or slower than  $t^{-\frac{1}{\gamma}}$ , the corresponding inhomogeneous model is weakly ergodic.

**Lemma 3.5.** *Suppose  $\hat{\varepsilon} : \mathbb{Z}_+ \rightarrow (0, \bar{\varepsilon}]$  is such that  $\hat{\varepsilon}(t) \rightarrow 0$  and  $\hat{\varepsilon}(t) = \Omega\left(t^{-\frac{1}{\gamma}}\right)$  as  $t \rightarrow \infty$ . Then the Markov chain induced by the inhomogeneous model of evolution is weakly ergodic.*

*Proof.* Let

$$\bar{z} \in \arg \min_{z' \in C} CR(\{z'\}) + \alpha(\{z'\})$$

be a limit point that achieves  $\gamma$ . Let  $\bar{\alpha} = \alpha(\{\bar{z}\})$ , and let  $\bar{C} \in \{C_1, \dots, C_M\}$  be such that  $\bar{z} \in \bar{C}$ . For all  $z \in S$ , let  $H_{z\bar{z}} = \{h : h \text{ is a path from } z \text{ to } \bar{z} \text{ and } \tilde{c}(h) = r(z, \bar{z})\}$

<sup>14</sup>Since  $C_m$  is aperiodic under  $P(\varepsilon)$ , there exists  $N$  such that  $P_{zz'}^{(n)}(\varepsilon) > 0$  for all  $z, z' \in C_m$  and  $n \geq N$ . Therefore,  $\alpha(\{z'\}) < \infty$  for all  $z' \in C_m$ .

<sup>15</sup>If  $\bar{P}_{zz'} > 0$ , then  $P_{zz'}(\varepsilon) = k_{zz'} \varepsilon^{c(z, z')} (1 + f_{zz'}(\varepsilon)) \rightarrow \bar{P}_{zz'} > 0$  requires  $c(z, z') = 0$ . Since  $C_m$  is irreducible under  $\bar{P}$ , for any  $z, z' \in C_m$ , there exist  $z_1, \dots, z_n$  such that  $\bar{P}_{zz_1} \bar{P}_{z_1 z_2} \cdots \bar{P}_{z_n z'} > 0$ . So, there exists a zero cost path from  $z$  to  $z'$ .

be the set of minimum cost paths from  $z$  to  $\bar{z}$ . For each  $z \in S$ , fix  $h_{z\bar{z}} \in H_{z\bar{z}}$  and let  $d(z, \bar{z}) = \tilde{d}(h_{z\bar{z}})$  be the length of path  $h_{z\bar{z}}$ . Then, as  $\varepsilon \rightarrow 0$ ,

$$\begin{aligned} P_{z\bar{z}}^{(d(z,\bar{z}))}(\varepsilon) &= \text{probability of transitioning from } z \text{ to } \bar{z} \text{ in } d(z, \bar{z}) \text{ steps under } P(\varepsilon) \\ &= \Theta(\text{probability of transitioning from } z \text{ to } \bar{z} \text{ along } h_{z\bar{z}} \text{ under } P(\varepsilon)) \\ &= \Theta\left(\varepsilon^{r(z,\bar{z})}\right) \text{ by construction.} \end{aligned}$$

Let  $\hat{z} \in S$  be such that  $r(\hat{z}, \bar{z}) = CR(\{\bar{z}\})$ . Then  $r(\hat{z}, \bar{z}) \geq r(z, \bar{z})$  for all  $z \in S$ . Since

$$P_{z\bar{z}}^{(d(z,\bar{z}))}(\varepsilon) = \Theta\left(\varepsilon^{r(z,\bar{z})}\right) \quad \text{and} \quad P_{\hat{z}\bar{z}}^{(d(\hat{z},\bar{z}))}(\varepsilon) = \Theta\left(\varepsilon^{r(\hat{z},\bar{z})}\right),$$

we have

$$P_{z\bar{z}}^{(d(z,\bar{z}))}(\varepsilon) = \Omega\left(P_{\hat{z}\bar{z}}^{(d(\hat{z},\bar{z}))}(\varepsilon)\right) = \Omega\left(\varepsilon^{CR(\{\bar{z}\})}\right).$$

Let  $N \in \arg \min_{n \geq 1} \{\max_{z \in \bar{C}} r(z, \bar{z}, n)\}$ . Then, for all  $n \geq N$ ,

$$P_{\bar{z}\bar{z}}^{(n)}(\varepsilon) \geq \sum_{z \in \bar{C}} P_{\bar{z}z}^{(n-N)}(\varepsilon) P_{z\bar{z}}^{(N)}(\varepsilon) = \Omega\left(\varepsilon^{\bar{\alpha}}\right).$$

Let  $L = N + \max_{z \in S} d(z, \bar{z})$ . For each  $k \in \mathbb{Z}_+$ , let  $n_k \in \{kL, \dots, kL + (L-1)\}$  be such that  $\hat{\varepsilon}(n_k) = \min \{\hat{\varepsilon}(t) : kL \leq t \leq kL + (L-1)\}$ . Since  $\hat{\varepsilon}(t) = \Omega\left(t^{-\frac{1}{\gamma}}\right)$  as  $t \rightarrow \infty$ , there exists constant  $K > 0$  such that for all sufficiently large  $k$ ,

$$\hat{\varepsilon}(n_k) \geq \min \left\{ K(kL)^{-\frac{1}{\gamma}}, \dots, K(kL + (L-1))^{-\frac{1}{\gamma}} \right\}.$$

So,  $\hat{\varepsilon}(n_k) = \Omega\left(k^{-\frac{1}{\gamma}}\right)$  as  $k \rightarrow \infty$ .

Suppressing the possible dependence of the minimizers on  $k$  in the notation, let  $x$  and  $x'$  be the minimizers of

$$\min_{z, z' \in S} \sum_{z'' \in S} \min \left\{ P_{zz''}^{(kL, kL+L)}, P_{z'z''}^{(kL, kL+L)} \right\}.$$

As  $k \rightarrow \infty$ ,

$$\begin{aligned} &1 - \tau_1 \left( P^{(kL, kL+L)} \right) \\ &= \min_{z, z' \in S} \sum_{z'' \in S} \min \left\{ P_{zz''}^{(kL, kL+L)}, P_{z'z''}^{(kL, kL+L)} \right\} \\ &= \sum_{z'' \in S} \min \left\{ P_{xz''}^{(kL, kL+L)}, P_{x'z''}^{(kL, kL+L)} \right\} \\ &\geq \min \left\{ P_{x\bar{z}}^{(kL, kL+L)}, P_{x'\bar{z}}^{(kL, kL+L)} \right\} \\ &\geq \min \left\{ P_{x\bar{z}}^{(kL, kL+d(x,\bar{z}))} P_{\bar{z}\bar{z}}^{(kL+d(x,\bar{z}), kL+L)}, P_{x'\bar{z}}^{(kL, kL+d(x',\bar{z}))} P_{\bar{z}\bar{z}}^{(kL+d(x',\bar{z}), kL+L)} \right\} \\ &= \min \left\{ \Omega \left( P_{\hat{z}\bar{z}}^{(d(\hat{z},\bar{z}))}(\hat{\varepsilon}(n_k)) \right) \Omega \left( \hat{\varepsilon}(n_k)^{\bar{\alpha}} \right), \Omega \left( P_{\hat{z}\bar{z}}^{(d(\hat{z},\bar{z}))}(\hat{\varepsilon}(n_k)) \right) \Omega \left( \hat{\varepsilon}(n_k)^{\bar{\alpha}} \right) \right\} \\ &= \Omega \left( \hat{\varepsilon}(n_k)^{CR(\{\bar{z}\}) + \bar{\alpha}} \right) \\ &= \Omega \left( \hat{\varepsilon}(n_k)^\gamma \right) = \Omega \left( \frac{1}{k} \right). \end{aligned}$$

Let  $t_k = kL$  for each  $k \in \mathbb{Z}_+$ . Then

$$\sum_{k=0}^{\infty} \left(1 - \tau_1 \left(P^{(t_k, t_{k+1})}\right)\right) = \sum_{k=0}^{\infty} \left(1 - \tau_1 \left(P^{(kL, kL+L)}\right)\right) = \infty.$$

Therefore, the inhomogeneous chain defined by  $\{P(\hat{\varepsilon}(t)) : t \in \mathbb{Z}_+\}$  is weakly ergodic by Doeblin's theorem.  $\square$

Next, we verify that Lemma 3.5 provides a correct bound in Example 3.4.

**Example 3.6.** Continuing example 3.4, we have

$$\gamma = \min_{z \in C} CR(\{z\}) + \alpha(\{z\}) = CR(\{AA\}) + \alpha(\{AA\}) = 3.$$

Let  $\hat{\varepsilon}(t) = (t+2)^{-\frac{1}{3}}$  so that  $\hat{\varepsilon}(t) = \Omega\left(t^{-\frac{1}{\gamma}}\right)$ , and let  $t_k = k$ . Then

$$\min_{z, z' \in S} \sum_{z'' \in S} \min \left\{ P_{zz''}^{(k, k+1)}, P_{z'z''}^{(k, k+1)} \right\} = \Theta(\hat{\varepsilon}(k)^2) = \Theta\left(\frac{1}{k^{\frac{2}{3}}}\right).$$

Therefore,

$$\sum_{k=0}^{\infty} \left(1 - \tau_1 \left(P^{(t_k, t_{k+1})}\right)\right) = \infty,$$

and the chain is weakly ergodic by Doeblin's theorem.  $\square$

As expected, when the minimum coradius set is aperiodic, then its coradius gives the correct bound on the required rate of decline.

**Corollary 3.7.** *Let  $\bar{C} \in \arg \min_{C_m \in \{C_1, \dots, C_M\}} CR(C_m)$ , and let  $\beta = CR(\bar{C})$ . Suppose  $\bar{C}$  is aperiodic under  $\bar{P}$ , and  $\hat{\varepsilon} : \mathbb{Z}_+ \rightarrow (0, \bar{\varepsilon}]$  is such that  $\hat{\varepsilon}(t) \rightarrow 0$  and  $\hat{\varepsilon}(t) = \Omega\left(t^{-\frac{1}{\beta}}\right)$  as  $t \rightarrow \infty$ . Then the Markov chain induced by the inhomogeneous model of evolution is weakly ergodic.*

*Proof.* Since  $\bar{C}$  is aperiodic and irreducible under  $\bar{P}$ , there exists  $N$  such  $\bar{P}_{zz'}^{(n)} > 0$  for all  $z, z' \in \bar{C}$  and  $n \geq N$ . Therefore, for all  $z' \in \bar{C}$ ,

$$\alpha(\{z'\}) = \min_{n \geq 1} \max_{z \in \bar{C}} r(z, z', n) = 0.$$

$\square$

### 3.3 Strong Ergodicity

Having established the conditions guaranteeing weak ergodicity, we now turn to strong ergodicity. Sufficient conditions for strong ergodicity are derived by appealing to the following finite state space version of Isaacson and Madsen's theorem.

**Theorem 3.8** (Isaacson and Madsen).<sup>16</sup> *A weakly ergodic inhomogeneous Markov chain on finite state space  $S$  is strongly ergodic if for all  $t$  there exists an invariant distribution  $\pi^{(t)}$  of  $P^{(t,t+1)}$  such that*

$$\sum_{t=0}^{\infty} \sum_{z \in S} |\pi_z^{(t)} - \pi_z^{(t+1)}| < \infty.$$

Moreover, if  $\pi = \lim_{t \rightarrow \infty} \pi^{(t)}$ , then for all  $t \in \mathbb{Z}_+$  and  $z, z' \in S$ ,

$$\lim_{t' \rightarrow \infty} P_{zz'}^{(t,t')} = \pi_{z'}.$$

The following theorem shows that if an additional restriction on the variation of the mutation parameter is added to the conditions guaranteeing weak ergodicity, then strong ergodicity is obtained. Moreover, the limiting distribution of the inhomogeneous model is precisely the stochastically stable distribution of the homogeneous model.

**Theorem 3.9.**<sup>17</sup> *Suppose  $\hat{\varepsilon} : \mathbb{Z}_+ \rightarrow (0, \bar{\varepsilon}]$  satisfies the conditions of Lemma 3.5. Suppose further that  $\sum_{t=1}^{\infty} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^+ < \infty$ . Then the Markov chain induced by the inhomogeneous model of evolution is strongly ergodic. Moreover, for all  $t \in \mathbb{Z}_+$  and  $z, z' \in S$ ,*

$$\lim_{t' \rightarrow \infty} P_{zz'}^{(t,t')} = \pi_{z'}^*,$$

where  $\pi^*$  is the limiting distribution given in (1).

*Proof.* The weak ergodicity of the inhomogeneous chain  $\{P(\hat{\varepsilon}(t)) : t \in \mathbb{Z}_+\}$  follows from Lemma 3.5. Let  $\pi^{(t)}$  be the unique invariant distribution for the homogeneous chain  $P(\hat{\varepsilon}(t))$ . We will show that  $\sum_{t=0}^{\infty} \sum_{z \in S} |\pi_z^{(t)} - \pi_z^{(t+1)}| < \infty$ . The strong ergodicity then follows from Isaacson and Madsen's theorem.

For all  $z \in S$ , let  $H_z = \{h : h \text{ is a } z\text{-tree}\}$ .<sup>18</sup> For each  $\varepsilon \in (0, \bar{\varepsilon}]$ , let  $\pi(\varepsilon)$  denote the unique invariant distribution for the homogeneous chain defined by  $P(\varepsilon)$ . By the familiar tree-surgery theorem,<sup>19</sup>

$$\pi_z(\varepsilon) = \frac{q_z(\varepsilon)}{\sum_{z' \in S} q_{z'}(\varepsilon)},$$

where

$$q_z(\varepsilon) = \sum_{h \in H_z} \prod_{(z', z'') \in h} P_{z'z''}(\varepsilon).$$

For any  $z$ -tree  $h$ , let  $\tilde{c}(h) = \sum_{(z', z'') \in h} c(z', z'')$ . Then

$$\prod_{(z', z'') \in h} P_{z'z''}(\varepsilon) = \prod_{(z', z'') \in h} \left( k_{z'z''} \varepsilon^{c(z', z'')} (1 + f_{z'z''}(\varepsilon)) \right) = k_h \varepsilon^{\tilde{c}(h)} (1 + f_h(\varepsilon))$$

<sup>16</sup>See, Isaacson and Madsen [7] Theorem V4.3.

<sup>17</sup>The earlier version of the theorem required  $\sum_{t=1}^{\infty} |\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t)| < \infty$ . I would like to thank a referee for suggesting that this condition can be weakened to  $\sum_{t=1}^{\infty} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^+ < \infty$ .

<sup>18</sup>A  $z$ -tree is a directed graph on  $S$  such that every  $z' \in S \setminus \{z\}$  has exactly one successor and there are no closed loops.

<sup>19</sup>See, for example, Freidlin and Wentzel [5].

for some constant  $k_h > 0$  and some  $f_h(\varepsilon)$  that is  $o(1)$  as  $\varepsilon \rightarrow 0$ . Moreover, since  $c(z', z'') \in \mathbb{Z}_+ \cup \{+\infty\}$  and  $f_{z'z''}(\varepsilon)$  is Lipschitz continuous on  $[0, \bar{\varepsilon}]$  for all  $z', z'' \in S$ ,  $f_h(\varepsilon)$  is Lipschitz continuous on  $[0, \bar{\varepsilon}]$  as well. Let  $\delta_z = \min_{h \in H_z} \tilde{c}(h)$ . Since  $P(\varepsilon)$  is irreducible,  $\delta_z < \infty$  for all  $z \in S$ . Let  $\hat{\delta} = \min_{z \in S} \delta_z$ . Then

$$q_z(\varepsilon) = \sum_{h \in H_z} k_h \varepsilon^{\tilde{c}(h)} (1 + f_h(\varepsilon)) = K_z \varepsilon^{\delta_z} (1 + f_z(\varepsilon)),$$

where  $K_z > 0$  and  $f_z(\varepsilon)$  is Lipschitz continuous on  $[0, \bar{\varepsilon}]$  and  $o(1)$ . This in turn implies

$$\sum_{z' \in S} q_{z'}(\varepsilon) = K \varepsilon^{\hat{\delta}} (1 + f(\varepsilon))$$

where  $K > 0$  and  $f(\varepsilon)$  is Lipschitz continuous on  $[0, \bar{\varepsilon}]$  and  $o(1)$ . Therefore,

$$\pi_z(\varepsilon) = \frac{q_z(\varepsilon)}{\sum_{z' \in S} q_{z'}(\varepsilon)} = \left( \frac{K_z \varepsilon^{\delta_z}}{K \varepsilon^{\hat{\delta}}} \right) \left( \frac{1 + f_z(\varepsilon)}{1 + f(\varepsilon)} \right).$$

Since  $\delta_z - \hat{\delta} \in \mathbb{Z}_+$  and  $f(\varepsilon)$  is  $o(1)$ ,  $\pi_z(\varepsilon)$  is Lipschitz continuous on  $[0, \bar{\varepsilon}]$ . So, there exists  $M_z < \infty$  such that  $|\pi_z(\varepsilon') - \pi_z(\varepsilon)| < M_z |\varepsilon' - \varepsilon|$  for all  $\varepsilon, \varepsilon' \in [0, \bar{\varepsilon}]$ .

Next, for all  $n \geq 1$ ,

$$\hat{\varepsilon}(n) - \hat{\varepsilon}(0) = \sum_{t=0}^{n-1} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^+ - \sum_{t=0}^{n-1} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^-.$$

Since  $\hat{\varepsilon}(n) > 0$  for all  $n$ ,

$$\sum_{t=0}^{n-1} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^- < \hat{\varepsilon}(0) + \sum_{t=0}^{n-1} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^+.$$

So,

$$\sum_{t=0}^{\infty} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^- < \infty,$$

and

$$\sum_{t=0}^{\infty} |\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t)| = \sum_{t=0}^{\infty} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^+ + \sum_{t=0}^{\infty} (\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t))^- < \infty.$$

Therefore,

$$\begin{aligned} \sum_{t=0}^{\infty} \sum_{z \in S} |\pi_z^{(t)} - \pi_z^{(t+1)}| &= \sum_{z \in S} \sum_{t=0}^{\infty} |\pi_z(\hat{\varepsilon}(t+1)) - \pi_z(\hat{\varepsilon}(t))| \\ &< \sum_{z \in S} \sum_{t=0}^{\infty} M_z |\hat{\varepsilon}(t+1) - \hat{\varepsilon}(t)| \\ &< \infty \end{aligned}$$

as desired.  $\square$

An immediate consequence of this theorem is that for any mutation parameter that goes to zero monotonically at the rate of  $t^{-\frac{1}{\gamma}}$  or slower, the resulting inhomogeneous model of evolution is strongly ergodic.



## 4 Concluding Remarks

This paper considered time-dependent mutations in a class of Markov models that includes those with periodic cycles. It has shown that as long as the mutation parameter converges to zero slowly enough and its variation is finite, the resulting time-inhomogeneous model has a limiting distribution regardless of the details of the mutation process. However, since the bound on the required rate is found to be  $t^{-\frac{1}{\gamma}}$ , the result suggests that the mutation may need to decrease very slowly if  $\gamma$  is large.

## References

- [1] Bergin, J. and B. Lipman (1996). Evolution with State-Dependent Mutations. *Econometrica*, 64: 943-56.
- [2] Chen, H. and Y. Chow (2001). On the Convergence of Evolution Processes with Time-Varying Mutations and Local Interaction. *Journal of Applied Probability*, 38: 301-323.
- [3] Doeblin, W. (1937). Le Cas Discontinu des Probabilites en Chaine. *Publ. Fac. Sci. Univ. Masaryk*, 236: 3-13, 592 (a correction).
- [4] Ellison, G. (2000). Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution. *Review of Economic Studies*, 67: 17-45.
- [5] Freidlin, M. and A. Wentzel (1984). *Random Perturbation of Dynamical Systems*. New York: Springer Verlag.
- [6] Grimmett, G. R. and D. R. Stirzaker. (1992). *Probability and Random Processes*. Oxford: Oxford University Press.
- [7] Isaacson, D. and R. Madsen. (1976). *Markov Chains: Theory and Application*. New York: Wiley.
- [8] Kandori, M., G. Mailath, and R. Rob. (1993). Learning to Play Equilibria in Games with Stochastic Perturbations. *Econometrica*, 61: 29-56.
- [9] Robles, J. (1998). Evolution with Changing Mutation Rates. *Journal of Economic Theory*, 79: 207-223.
- [10] Sandholm, W. and A. Pauzner. (1998). Evolution, Population Growth, and History Dependence. *Games and Economic Behavior*, 22: 84-120.
- [11] Seneta, E. (1981). *Non-negative Matrices and Markov Chains*. New York: Springer-Verlag.
- [12] Young, P. (1993). The Evolution of Conventions. *Econometrica*, 61: 57-83.