Markov chains with self-interactions

P. Del Moral* L. Miclo*

March 14, 2002

Abstract

In this article we study a class of time self-interacting "Markov" chain models. We propose a novel theoretical basis based on measure valued processes and semigroup technics to analyze their asymptotic behavior as the time parameter tends to infinity. We exhibit different types of decays to equilibrium depending on the level of interaction. We illustrate these results in a variety of examples including Gaussian or Poisson self-interacting models. We analyze the long time behavior of a new class of evolutionary self-interacting chain models. These genetic type algorithms can also be regarded as reinforced stochastic explorations of an environment with obstacles related to a potential function.

1 Introduction

A (time) self-interacting Markov chain model (abbreviate **SIMC**) is a collection of random variables with a countable time index. In contrast to traditional Markov chains their evolution in time may depend on the occupation measure of the past values.
This form of interaction can be interpretated in various ways. In biology this structure is used to model some population dynamical structures [3, 12, 14, 21] as well as polymers and tree evolutions [13, 18]. This interaction structure can also express the reinforcement degree of edges or vertices in a graph visited by a random walk [15]. The question of recurrence and transience of the latter has been initiated in [3, 15]. During the last two decades this subject has been investigated in various articles [1, 10, 16, 17, 19, 22]. This is still an active research area, to our knowledge we don't know for instance if the reinforced random walk on the integer lattice $\mathbb{Z}^2$ is recurrent or not.

The SIMC models presented in this article can also be related to the continuous time self-interacting diffusions analyzed in [2]. In this work the authors discuss a gradient type diffusion on a compact Riemannian manifold with linear self-interactions. They connect the shape properties of the drift function with the nature of the long time behavior of the occupation measure of the diffusion. The strategy consists in proving that the set of all possible limiting measures is the attractor free set of a deterministic dynamical semi-flow in the set of bounded measures on the manifold.

*Laboratoire de Statistiques et Probabilités, UMR 5583, CNRS and Université Paul Sabatier, 31062 Toulouse cedex 4, France, delmoral(miclo)@cict.fr

Although our investigations have been partly influenced by the latter referenced papers this article is not concerned with the same class of self-interacting processes and it doesn't discuss the same stability questions. To our knowledge the questions of the rates of convergence and their precise connections with the strength of interaction has not be covered in the literature on self-interacting processes. The approach we have taken here rather comes from a different background. In previous works [4, 5, 6] we proposed an interacting particle technique for solving an abstract class of measure valued processes associated to a given abstract mapping Φ on distribution space. Under appropriate regularity conditions we proved that the empirical measures of the particle systems converge as the size and time parameters tend to infinity to the fixed point of Φ (whenever it exists). In the context of Feynman-Kac type distribution evolutions these limiting distributions can be used to analyze the limiting behavior of a killed particle evolving in an environment with obstacles related to a potential function. In this situation the numerical solving of these limiting distributions allows to compute the Lyapunov exponent of a class of Feynman-Kac-Schrödinger semigroup [8].
In this article we propose an alternative interpretation of these limiting distributions. We associate to an abstract mapping Φ a class of SIMC models and we propose a set of regularity conditions underwhich the resulting occupation measures converge as the time tends to infinity to the desired equilibrium distribution. We analyze in this framework the evolutionary SIMC versions of the genetic type particle models studied in [5, 6, 8, 9]. We also discuss other types of interactions including the SIMC versions of Gaussian and Poisson mean field particle interactions. Together with the modeling of these SIMC we provide a precise analysis of the asymptotic behavior of their occupation measures. We exhibit different types of decays to equilibrium in terms of the level of interaction in the models. We will also discuss the long time behavior of a model of ϵ-interacting random variables for which these different decays are sharp. Unless we make some supplementary hypothesis on the regularity of Φ this example indicates that the estimates we obtained cannot be improved.

To our knowledge the abstract class of SIMC models presented in this article has not been covered in the literature. The precise connections between the interaction structure and the decays to equilibrium also seems to be the first result of this kind. The evolutionary and genetic type SIMC models presented in this article can also be regarded as novel reinforced stochastic exploration model of an environment with obstacles related to a potential function.
The self-interacting versions of the Moran type particle models in continuous time presented in [7, 8] will lead to a new class of models with self-interacting jumps. The analysis of these self-interacting evolutionary processes is under study and it will hopefully be discussed in a forthcoming article.

1.1 Description of the models and main results

Let X be a stochastic process with discrete time index $n \in \mathbb{N}$ and taking values in some measurable space $(E, \mathcal{E})$. We suppose it is defined on some probability space $(\Omega, F, \mathbb{P})$ with a nondecreasing family of σ-field F_n, $n \geq 0$. The adapted process X is called an SIMC when the state X_n at time n depends on the occupation measure of the previous states $X_0, \ldots, X_{n-1}$. For instance suppose we are given a mapping Φ from the set of probability measures $\mathcal{P}(E)$ on E into itself. We can associate to the latter the SIMC model defined by the transitions

$$\mathbb{P}(X_n \in dx \mid F_{n-1}) = \Phi\left(\frac{1}{n}\sum_{p=0}^{n-1} \delta_{X_p}\right)(dx) \tag{1}$$

(to get rid of measurability problems, we will always implicitly assume that $\Phi : \mathcal{P}(E) \to \mathcal{P}(E)$ is measurable, where $\mathcal{P}(E)$ is endowed with the σ-algebra generated by the evaluation mappings $\mathcal{P}(E) \ni m \mapsto m[f]$, for f bounded and measurable).
To illustrate this abstract class of SIMC model we already present an example which was at the origin of our investigations. Let g be a positive and measurable function and let K be a Markov transition on E. We associate to the pair (g, K) a non-linear mapping Φ on $\mathcal{P}(E)$ defined for any $m \in \mathcal{P}(E)$ and for any bounded measurable test function f on E as follows

$$\Phi(m)(f) = \frac{m(g\ (Kf))}{m(g)} \tag{2}$$

with the traditional notations

$$Kf(x) = \int_E K(x, dy)\ f(y) \quad \text{and} \quad m(f) = \int_E f(x)\ m(dx)$$

In this situation transition (1) reads

$$\Phi\left(\frac{1}{n}\sum_{p=0}^{n-1} \delta_{X_p}\right)(dx) = \sum_{p=0}^{n-1} \frac{g(X_p)}{\sum_{q=0}^{n-1} g(X_q)}\ K(X_p, dx) \tag{3}$$

The resulting SIMC model can be regarded as a genetic type algorithm. More precisely we readily see from (3) that the particle X_n first selects a site X_p, $0 \leq p < n$, with a probability proportional to its fitness $g(X_p)$, $0 \leq p < n$. Then it performs an elementary move according to the transition K.

This type of genetic self-interactions arise in many human endeavors. For instance in visiting a city E some sites are more attractive than others. The fitness or potential function g models the attraction level of the city's areas. The Markov transition K models the way the person explores randomly each place. In exploring the city with the above two step selection/mutation mechanism the person also tends to be attracted by familiar places which have been visited several times. This genetic SIMC model is close to the reinforced random walk model proposed in [3] but the non-linear structure of interaction of these models differ. In particular and to our knowledge none of the reinforced random walk models presented in the literature are related to a (non-constant) potential or fitness function but only keep track of the number of time an edge or a vertex has been visited.

The long time behavior of traditional reinforced random walks is also generally studied on finite graphs or on the integer lattice using urn's processes and random environment techniques or Robbins-Monro's type approximation analysis. As mentioned in the introduction we will not discuss the same stability questions and our strategy is rather based on measure valued processes and semigroup techniques. We also make no restrictions on the state space and we illustrate most of our results on a class of genetic type SIMC models. Other examples including the SIMC versions of Gaussian and Poisson mean field interactions will also be considered.

It is transparent from (1) that the sequence of occupation measures

$$S_n = \frac{1}{n+1} \sum_{p=0}^{n} \delta_{X_p}$$

forms a time inhomogeneous Markov process in distribution space. More precisely for any bounded measurable function G on $\mathcal{P}(E)$ we have

$$\mathbb{E}(G((n+1)S_n) \mid S_{n-1}) = \int_E G(n\ S_{n-1} + \ \delta_x)\ \Phi(S_{n-1})(dx)$$

When the mapping Φ is regular enough it is natural to expect that S_n $\mathcal{F}$-weakly converges as the time parameter n tends to infinity to the fixed point $\mu = \Phi(\mu)$ (whenever it exists) of the mapping Φ.

A set $\mathcal{F}$ of $\mathcal{E}$-measurable functions will be called a test functions collection if for all $f \in \mathcal{F}$, $\|f\|_\infty \leq 1$ and if $d_\mathcal{F}$ is a complete metric on $\mathcal{P}(E)$, where by definition

$$\forall\ p_1, p_2, \in \mathcal{P}(E), \qquad d_\mathcal{F}(p_1, p_2) \ = \ \sup_{f \in \mathcal{F}} |p_1(f) - p_2(f)|$$

The first example of a test functions collection one think about is the largest possible choice

$$\mathcal{F} \ = \ \{f\ \mathcal{E} - \text{measurable} \ : \ \|f\|_\infty \leq 1\}$$

Then the distance $d_\mathcal{F}$ is complete, since it is given by twice the total variation norm.
One also recover the latter by considering

$$\mathcal{F} \ = \ \{f/\|f\|_\infty \ : \ f \in C_b(E) \setminus \{0\}\}$$

where $C_b(E)$ is the set of bounded continuous functions, if E is a Polish topological space (endowed with its Borelian σ-field $\mathcal{E}$). But in this context, one can also end up with a distance $d_\mathcal{F}$ metrizing the weak convergence, by considering for instance

$$\mathcal{F} \ = \ \{f_n/(M_n\, \|f_n\|_\infty) \ : \ n \in \mathbb{N}\}$$

where $(f_n)_{n\in\mathbb{N}}$ is weak convergence determining sequence of $C_b(E)$ (recall that this means that a sequence of probability measures $(m_p)_{p\in\mathbb{N}}$ is weakly convergent if and only if for all $n \in \mathbb{N}$, $(m_p(f_n))_{p\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$) and where $(M_n)_{n\in\mathbb{N}}$ is a sequence of positive real numbers diverging to infinity.
We will use the following regularity condition.

(HΦ) *There exists a test functions collection $\mathcal{F}$ such that for any $(f, \mu) \in$*

$\mathcal{F} \times \mathcal{P}(E)$ we can associate a subset $H(f,\mu) \subset \mathcal{F}$ and a collection of numbers $b_{f,\mu}(h) \in [0,1)$, $h \in H(f,\mu)$, satisfying the following inequality

$$|\Phi(\mu)(f) - \Phi(\nu)(f)| \leq \sum_{h \in H(f,\mu)} b_{f,\mu}(h)\ |\mu(h) - \nu(h)|$$

for any $\nu \in \mathcal{P}(E)$ with $\beta_{\mathcal{F}}(\Phi) = \sup\{\sum_{h\in H(f,\mu)} b_{f,\mu}(h)\ ;\ f \in \mathcal{F},\ \mu \in \mathcal{P}(E)\} \in [0,1)$.

If we denote by $\beta(K)$ the contraction coefficient of the total variation distance of distributions induced by the integral Markov operator K then we will check that the genetic type self interacting chain (3) satisfies this condition with the class of functions f, $\|f\|_\infty \leq 1$, with oscillations $\operatorname{osc}(f) \leq 1$ and

$$\beta_{\mathcal{F}}(\Phi) = \frac{1}{g_{min}}(\operatorname{osc}(g) + \|g\|_\infty)\ \beta(K) < 1 \tag{4}$$

as soon as $g_{min} = \inf_E g > 0$ and

$$\sup_{x,y} \frac{g(x)}{g(y)} < \frac{1}{2}\ \left(\frac{1}{\beta(K)} + 1\right)$$

The case $g = 1$ correspond to linear interaction mappings $\Phi(m) = mK$. In this situation the contraction parameter $\beta_{\mathcal{F}}(\Phi) = \beta(K)$ coincide with the contraction coefficient of K. In this special case the chain is not attracted by sites with high g-fitness but it has simply more chance to visit sites which have been visited several times. Our first main result can be stated as follows.

Theorem 1.1 *Suppose the mapping Φ satisfies condition* **(HΦ)** *for some set of functions $\mathcal{F}$. Then there exists a unique fixed point $\mu = \Phi(\mu) \in \mathcal{P}(E)$ and the occupation measure S_n $\mathcal{F}$-weakly converges to μ in probability.. In addition for any $f \in \mathcal{F}$ we have three different types of $\mathbb{L}_2$-mean error decays:*

$$\begin{array}{lcl}
\beta_{\mathcal{F}}(\Phi) > 1/2 & \Rightarrow & n^{2(1-\beta_{\mathcal{F}}(\Phi))}\ \mathbb{E}((S_n(f) - \mu(f))^2) \leq c \\
\beta_{\mathcal{F}}(\Phi) < 1/2 & \Rightarrow & (1 - 2\beta_{\mathcal{F}}(\Phi))\ n\ \mathbb{E}((S_n(f) - \mu(f))^2) \leq c \\
\beta_{\mathcal{F}}(\Phi) = 1/2 & \Rightarrow & n/\log(n)\ \mathbb{E}((S_n(f) - \mu(f))^2) \leq c
\end{array}$$

for some finite constant $c < \infty$ which doesn't depend on the function f nor on the time parameter.

When the Markov transition K is given by

$$K(x,dy) = \epsilon\ \delta_x(dy) + (1-\epsilon)\ \mu(dy)$$

for some $\epsilon \in (0,1]$ and some measure $\mu \in \mathcal{P}(E)$ the SIMC model produces with probability $(1-\epsilon)$ independent and identically distributed random variables with distribution μ and with a probability ϵ it tends to repeat the previous sites with a probability proportional to the number of times they have been visited. For this particular model with linear interactions most of the calculations can be done explicitly. For instance we have that $\beta_{\mathcal{F}}(\Phi) = \beta(K) = \epsilon$. In section 2.1.1 we propose a detailed analysis of this sequence of ϵ-interacting variables. We will show that in this situation the above three different behaviors are sharp.

So at least under the assumptions of theorem 1.1, the SIMC algorithm fulfills the goal for which it was introduced, namely to find the invariant probability μ for Φ. But in some cases, we can consider another type of SIMC to reach this objective. This will also enable us to come a little closer to usual simple reinforced walk, where the next step of the chain also depends specifically on its previous state.
We now suppose the mapping Φ can be written in the following form

$$\Phi(m) = mK_m \tag{5}$$

for some collection of Markov transitions K_m indexed by $m \in \mathcal{P}(E)$ (more precisely, a natural measurability property of K_m in m will always be implicitly assumed). In this situation an alternative model is defined by replacing (1) by the elementary transitions

$$\mathbb{P}(X_n \in dx \mid F_{n-1}) = K_{\frac{1}{n}\sum_{p=0}^{n-1}\delta_{X_p}}(X_{n-1}, dx) \tag{6}$$

The study of this class of SIMC models is a little more involved than the previous one. Note that in this situation the pair (X_n, S_n) forms a Markov chain. For any bounded measurable functions G on the product space $(E \times \mathcal{P}(E))$ we have

$$\mathbb{E}(G(X_n, (n+1)S_n) \mid (X_{n-1}, S_{n-1})) = \int_E G(x, n\ S_{n-1} + \ \delta_x)\ K_{S_{n-1}}(X_{n-1}, dx)$$

To illustrate this new class of SIMC models let us present a couple of examples connected to the genetic search algorithm (2).
When the potential function g is strictly greater than 1 we can rewrite the genetic mapping Φ defined in (2) in the form (5). As noticed in [6] the choice of K_m is not unique. We can choose for instance

$$K_m(x, dz) \ = \ \int_E L_m(x, dy)\ K(y, dz) \tag{7}$$

where $L_m(x, dy)$ denotes the Markov transition defined by

$$L_m(x, dy) = 1/m(g)\ \delta_x(dy) + (1 - 1/m(g))\ \Psi'(m)(dy)$$

with

$$\Psi'(m)(f) = m(f\ g')/m(g') \quad \text{and} \quad g' = g - 1$$

To have an intuitive feel of the corresponding SIMC model we notice that the former transition is again decomposed in a two step selection/mutation evolution. The mutation mechanism of the two models coincide but in the former selection the particle X_{n+1} has a probability $1/S_n(g)$ for staying in the same place and a probability $1 - 1/S_n(g)$ for selecting randomly a site X_p, $p \leq n$, with distribution

$$\Psi'(S_n) = \sum_{p=0}^{n} \frac{g'(X_p)}{\sum_{p=0}^{n} g'(X_q)}\ \delta_{X_p}$$

Using the same notations as above suppose now the Markov transitions K_m are defined as follows

$$K_m(x, dy) \ = \ a\ K(x, dy) \ + \ b\ \Psi(m)(dy) \quad \text{with} \quad \Psi(m)(f) = m(fg)/m(g) \tag{8}$$

for some $a, b \in (0, 1)$ such that $a + b = 1$. Here the particle X_n decides randomly to perform a selection or a mutation transition. With a probability a it evolves according to the mutation transition K and with a probability b it selects a site X_p, $0 \leq p < n$, according to the discrete distribution

$$\Psi(\frac{1}{n}\sum_{p=0}^{n-1}\delta_{X_p}) = \sum_{p=0}^{n-1}\frac{g(X_p)}{\sum_{q=0}^{n-1} g(X_q)}\ \delta_{X_p}$$

Here again, when the transitions K_m are sufficiently regular, it is natural to expect that the occupation measures S_n again converge to the fixed point $\mu = \Phi(\mu) = \mu K_\mu$.

The asymptotic behavior of the SIMC model (6) will be studied with the stronger hypothesis:

(HK) *For each $m \in \mathcal{P}(E)$ there exists a unique invariant measure $\pi(m) = \pi(m)K_m \in \mathcal{P}(E)$. There also exists a test functions collection $\mathcal{F}$ such that for any $(f,\mu) \in (\mathcal{F}\times\mathcal{P}(E))$ we can associate a subset $H(f,\mu) \subset \mathcal{F}$ and a collection of numbers $b_{f,\mu}(h) \in [0,1)$, $h \in H(f,\mu)$, satisfying the following inequality*

$$|\pi(\mu)(f) - \pi(\nu)(f)| \leq \sum_{h\in H(f,\mu)} b_{f,\mu}(h)\ |\mu(h) - \nu(h)| \tag{9}$$

for any $\nu \in \mathcal{P}(E)$ with $\beta_{\mathcal{F}}(\pi) = \sup\{\sum_{h\in H(f,\mu)} b_{f,\mu}(h)\ ;\ f \in \mathcal{F}\} \in (0,1)$. In addition, for each pair of measures $(m_1, m_2) \in \mathcal{P}(E)\times\mathcal{P}(E)$ and any pair of functions $(f_1, f_2) \in \mathcal{F}\times\mathcal{F}$ with $\pi(m_1)(f_1) = 0 = \pi(m_2)(f_2)$ we have

$$\|K^n_{m_1}(f_1) - K^n_{m_2}(f_2)\|_\infty \ \leq\ \epsilon(n)\ [\ \|f_1 - f_2\|_\infty + \|m_1 - m_2\|_{tv}\ ],\quad with \sum_{n\geq 1}\epsilon(n) < \infty \tag{10}$$

Theorem 1.2 *Suppose condition* **(HK)** *is met for some set of functions $\mathcal{F}$. Then for each $f \in \mathcal{F}$ and $p > 1$ we have that*

$$\sup_{n\geq 1}\sqrt{n}\ \mathbb{E}(|S_n(f) - S^\pi_n(f)|^p)^{1/p} < \infty \qquad with \quad S^\pi_n = \frac{1}{n+1}\sum_{q=0}^{n}\ \pi(S_q)$$

In addition the mapping π has a unique fixed point μ and for any $f \in \mathcal{F}$ we have three different types of $\mathbb{L}_2$-mean error decays:

$$\begin{array}{lcl}
\beta_{\mathcal{F}}(\pi) > 3/4 & \Rightarrow & n^{2(1-\beta_{\mathcal{F}}(\pi))}\ \mathbb{E}((S^\pi_n(f) - \mu(f))^2) \leq c \\
\beta_{\mathcal{F}}(\pi) < 3/4 & \Rightarrow & \sqrt{n}\ \mathbb{E}((S^\pi_n(f) - \mu(f))^2) \leq c \\
\beta_{\mathcal{F}}(\pi) = 3/4 & \Rightarrow & \sqrt{n}/\log(n)\ \mathbb{E}((S^\pi_n(f) - \mu(f))^2) \leq c
\end{array}$$

for some finite constant $c < \infty$ which doesn't depend on the function f nor on the time parameter.

We readily observe that

$$K_m(x,dy) = \Phi(m)(dy) \implies \pi(m) = \Phi(m) \quad \text{and} \quad (\textbf{(H}K\textbf{)} \iff \textbf{(H}\Phi\textbf{)})$$

In such a situation we also have that $\beta_{\mathcal{F}}(\pi) = \beta_{\mathcal{F}}(\Phi)$ and theorem 1.1 gives when $\beta_{\mathcal{F}}(\Phi) \leq 3/4$ better rates of decays to equilibrium than theorem 1.2. We believe that the latter can be improved and we conjecture that the rates presented in theorem 1.1 also hold for the SIMC model (6).

An apparent difficulty with the SIMC model (6) is that the state of the chain X_n at time n not only depends on the occupation measure S_{n-1} but also on the previous visited site X_{n-1}. To deal with this difficulty we have added an auxiliary regularity condition on the composite mappings K^n_m. This new condition is more difficult to check in practice. In the further development of section 3.2 we illustrate the regularity condition **(H**K**)** with several examples with a respective constant, linear and non-linear mapping $\pi : \mathcal{P}(E) \to \mathcal{P}(E)$.

We already mention that the two evolutionary SIMC models (7) and (8) satisfy condition **(H**K**)** with the class of functions f, $\|f\|_\infty \leq 1$, with oscillations $\mathrm{osc}(f) \leq 1$. For instance in the second case we have

$$\pi(m) = \Psi(m)K_a \quad \text{and} \quad \beta_{\mathcal{F}}(\pi) = \frac{\mathrm{osc}(g) + \|g\|_\infty}{g_{min}} \; \frac{b}{1-\beta(K)a} \tag{11}$$

with the resolvent kernel $K_a = b\sum_{n\geq 0} a^n K^n$. Note that $\beta_{\mathcal{F}}(\pi) < 1$ as soon as the pair (g, K) is chosen such that $g_{min} > 0$ and

$$\sup_{x,y} \frac{g(x)}{g(y)} < 1 + \frac{a}{2b}\; (1 - \beta(K))$$

1.2 Some terminology and preliminary results

We let $(E, \mathcal{E})$ be a general measurable space and we denote by $\mathcal{P}(E)$ the set of all probability measures on $(E, \mathcal{E})$ with the total variation norm

$$\|m_1 - m_2\|_{tv} := \sup_{A\in\mathcal{E}} |m_1(A) - m_2(A)|$$

We recall that a Markov transition $K(x, dz)$ on E generates two integral operations from the left and from the right. The first one acting on the set $\mathcal{B}_b(E)$ of bounded $\mathcal{E}$-measurable functions $f : E \to \mathbb{R}$ and the second one on probability measures $m \in \mathcal{P}(E)$

$$Kf(x) = \int_E K(x, dy)\, f(y) \qquad mK(A) = \int_E m(dx)\; K(x, A)$$

We also define inductively the step Markov transitions $K^n(x, dy)$

$$K^n(x, A) = \int\; K(x, dy)\; K^{n-1}(y, A)$$

for each $(x, A) \in E \times \mathcal{E}$ and $n \geq 0$. For $n = 0$ we use the convention $K^0(x, dy) = \delta_x(dy)$.
Given a function $f \in \mathcal{B}_b(E)$ we denote $\|f\|_\infty$ and f_{min} respectively the supremum norm and the infimum

$$\|f\|_\infty = \sup_{x\in E} |f(x)| \quad \text{ and } \quad f_{min} = \inf_{x\in E} f(x)$$

Next we list some class of functions that will be of use in the article. We denote $\mathrm{Osc}_1(E)$ the convex set of $\mathcal{E}$-measurable functions f with oscillations less than one, that is

$$\mathrm{osc}(f) = \sup\left\{|f(x) - f(y)|\; ;\; x, y \in E\right\} \leq 1$$

We also introduce the subset

$$\mathrm{Bosc}_1(E) = \{f \in \mathrm{Osc}_1(E) \; ; \; \|f\|_\infty \leq 1\}$$

When the state space is a metric space (E, d) we denote by $\mathrm{Lip}_1(E)$ the convex set of all bounded measurable functions f on E such that

$$\|f\|_{\infty,\mathrm{L}} = \|f\|_\infty \vee \|f\|_\mathrm{L} \leq 1$$

where

$$\|f\|_\mathrm{L} = \sup\{|f(x) - f(y)|/d(x, y) \; , \; x, y \in E \quad : \quad d(x, y) \neq 0\}$$

To clarify the presentation we will slight abuse notations and unless otherwise stated c denotes a constant whose values may vary from line to line but it does not depend on the time parameter nor on the test function we will consider.

The degree of contraction of the total variation distance of probability measures induced by K is defined by

$$\beta(K) = \sup_{m_1, m_2 \in \mathcal{P}(E)} \frac{\|m_1 K - m_2 K\|_{tv}}{\|m_1 - m_2\|_{tv}} = \sup_{x,y \in E} \|K(x, .) - K(y, .)\|_{tv} \tag{12}$$

It can also be defined in terms of the Dobrushin's ergodic coefficient

$$\alpha(K) = 1 - \beta(K) = \inf \sum_{i=1}^{m} \min\left(K(x, A_i), K(z, A_i)\right) \tag{13}$$

where the infimum is taken over all $x, z \in E$ and all resolutions of E into partitions $\{A_i \; ; \; 1 \leq i \leq m\}$ and $m \geq 1$ (see for instance [11]).

To study the asymptotic behavior of SIMC it is convenient to describe $\beta(K)$ as the contraction of oscillations of functions induced by K.

Lemma 1.3 *For any pair of probability measures (m_1, m_2) on E we have*

$$\|m_1 - m_2\|_{tv} = \sup\left\{|m_1(f) - m_2(f)| \; ; \; f \in \mathrm{Osc}_1(E)\right\}$$

For any Markov transition K on E the contraction coefficient $\beta(K)$ can alternatively be defined by

$$\beta(K) = \sup\left\{\mathrm{osc}(Kf) \; ; \; f \in \mathrm{Osc}_1(E)\right\} \tag{14}$$

Proof: To prove the first assertion we recall that the total variation distance between two probability measures m_1 and m_2 can alternatively be defined in terms of a Hahn-Jordan decomposition $m_1 - m_2 = m^+ - m^-$ with

$$\|m_1 - m_2\|_{tv} = m^+(E) = m^-(E)$$

From this observation we have for any $f \in \mathrm{Osc}_1(E)$

$$\begin{aligned}
|m_1(f) - m_2(f)| &= |\int f(x) \; m^+(dx) - \int f(y) \; m^-(dy)| \\
&= \|m_1 - m_2\|_{tv} \; | \int (f(x) - f(y)) \; \frac{m^+(dx)}{m^+(E)} \; \frac{m^-(dy)}{m^-(E)} \; | \\
&\leq \|m_1 - m_2\|_{tv}
\end{aligned}$$

By taking the supremum over all $f \in \mathrm{Osc}_1(E)$ we find that

$$\sup\{|m_1(f) - m_2(f)| \;;\; f \in \mathrm{Osc}_1(E)\} \leq \|m_1 - m_2\|_{tv}$$

The reverse inequality can be check easily by noting that the indicator functions 1_A of measurable sets $A \in \mathcal{E}$, belong to $\mathrm{Osc}_1(E)$. Using this representation we obtain

$$\begin{aligned}
\beta(K) &= \sup_{x,y\in E} \|K(x,.) - K(y,.)\|_{tv} \\
&= \sup_{x,y\in E} \sup\{|Kf(x) - K(f)(y)| \;;\; f \in \mathrm{Osc}_1(E)\} \\
&= \sup\{\sup_{x,y} |Kf(x) - K(f)(y)| \;;\; f \in \mathrm{Osc}_1(E)\}
\end{aligned}$$

and therefore

$$\beta(K) \;=\; \sup\{\mathrm{osc}(K(f)) \;;\; f \in \mathrm{Osc}_1(E)\}$$

and the proof of (14) is now completed. ■

For any $\delta \neq 0$ and $n \geq 1$ we define

$$d_\delta(n) = \frac{n^\delta - 1}{\delta} \geq 0 \quad \text{and} \quad d_0(n) = \log(n)(= \lim_{\delta\to 0} d_\delta(n))$$

For $\delta > 0$, respectively $\delta < 0$, we notice that

$$d_\delta(n) \leq n^\delta/\delta\;, \quad \text{resp.} \quad d_\delta(n) \leq 1/|\delta|$$

Because of the central importance in this article of the next technical lemma we have included a detailed proof. However its proof is somewhat technical and may be skipped at a first reading.

Lemma 1.4 *For any $1 \leq p \leq n$ we have*

$$\log\left(\frac{n+1}{p+1}\right) \leq \sum_{q=p+1}^{n} \frac{1}{q} \leq \log\left(\frac{n}{p}\right) \tag{15}$$

If $\epsilon \geq 0$ and if we put $s = \sum_{p\geq 1} 1/p^2 (< \infty)$ then we have that

$$e^{-s\epsilon^2}\;\left(\frac{n+1}{p+1}\right)^\epsilon \leq \prod_{q=p+1}^{n}\left(1 + \frac{\epsilon}{q}\right) \leq \left(\frac{n}{p}\right)^\epsilon \tag{16}$$

For $0 \leq \epsilon \leq 1$ we have that

$$\frac{1}{e}\;\left(\frac{p}{n}\right)^\epsilon \leq \prod_{q=p+1}^{n}\left(1 - \frac{\epsilon}{q}\right) \leq \left(\frac{p+1}{n+1}\right)^\epsilon \tag{17}$$

and for $-1 \leq \epsilon \leq 1$

$$\frac{1}{2}\;\frac{d_\epsilon(n)}{n^\epsilon} \leq \sum_{p=1}^{n} \frac{1}{p^{1+\epsilon}} \leq 1 + \frac{d_\epsilon(n)}{n^\epsilon} \tag{18}$$

Proof: The proof of (15) is simply based on the two integral estimates

$$\log\left(\frac{n+1}{p+1}\right) = \sum_{q=p+1}^{n} \int_{q}^{q+1} \frac{dt}{t} \leq \sum_{q=p+1}^{n} \frac{1}{q} \leq \sum_{q=p+1}^{n} \int_{q-1}^{q} \frac{dt}{t} = \log\left(\frac{n}{p}\right)$$

To prove (16) we recall that for any $x \geq 0$

$$\log x \leq x - 1 \tag{19}$$

This yields for any $\epsilon \geq 0$

$$\prod_{q=p+1}^{n} \left(1 + \frac{\epsilon}{q}\right) = \exp \sum_{q=p+1}^{n} \log\left(1 + \frac{\epsilon}{q}\right) \leq \exp \sum_{q=p+1}^{n} \frac{\epsilon}{q} \leq \left(\frac{n}{p}\right)^{\epsilon}$$

To prove the reverse estimate we use the fact that

$$x\ \log x \geq x - 1 \tag{20}$$

for any $x \geq 0$. From this inequality we also obtain for $x > 0$

$$\frac{\log(x+1)}{x} \geq 1 - \frac{x}{x+1}$$

If we take $x = \epsilon/q$ we conclude that

$$\log\left(1 + \frac{\epsilon}{q}\right) \geq \frac{\epsilon}{q} - \frac{\epsilon^2}{q\ (q+\epsilon)} \geq \frac{\epsilon}{q} - \frac{\epsilon^2}{q^2}$$

Therefore using (15) we conclude that

$$\begin{aligned}\prod_{q=p+1}^{n} \left(1 + \frac{\epsilon}{q}\right) &= \exp \sum_{q=p+1}^{n} \log\left(1 + \frac{\epsilon}{q}\right) \geq \exp \sum_{q=p+1}^{n} [\frac{\epsilon}{q} - \frac{\epsilon^2}{q^2}] \\ &\geq \exp[-s\epsilon^2 + \sum_{q=p+1}^{n} \frac{\epsilon}{q}] \geq e^{-s\epsilon^2}\ \left(\frac{p+1}{n+1}\right)^{\epsilon}\end{aligned}$$

To prove (17) we use again (20) to check first that for any $0 \leq x < 1$

$$(1-x)\log(1-x) \geq -x$$

If we take $x = \epsilon/q$ then for any $q \geq 2$ we have that

$$\log(1-\epsilon/q) \geq -\frac{\epsilon}{q}\ \frac{q}{q-\epsilon} = -\frac{\epsilon}{q-\epsilon} \geq -\frac{\epsilon}{q-1}$$

This together with (15) imply that for any $1 \leq p \leq n$

$$\begin{aligned}\prod_{q=p+1}^{n} \left(1 - \frac{\epsilon}{q}\right) &= \exp \sum_{q=p+1}^{n} \log\left(1 - \frac{\epsilon}{q}\right) \geq \exp[-\epsilon \sum_{q=p+1}^{n} \frac{1}{q-1}] \\ &\geq \exp[-\frac{\epsilon}{p} - \sum_{q=p+1}^{n-1} \frac{\epsilon}{q}] \geq \exp[-\epsilon - \sum_{q=p+1}^{n} \frac{\epsilon}{q}] \geq e^{-1}\ \left(\frac{p}{n}\right)^{\epsilon}\end{aligned}$$

To prove the reverse estimation we again use (19) to check by (15) that

$$\prod_{q=p+1}^{n} \left(1 - \frac{\epsilon}{q}\right) = \exp \sum_{q=p+1}^{n} \log\left(1 - \frac{\epsilon}{q}\right) \leq \exp[-\epsilon \sum_{q=p+1}^{n} \frac{1}{q}] \leq \left(\frac{p+1}{n+1}\right)^{\epsilon}$$

To prove (18) we first suppose that $0 < \epsilon \leq 1$. In this case we have the estimates

$$\sum_{p=1}^{n} \frac{\epsilon}{p^{1+\epsilon}} \quad \leq \quad \epsilon\ (1 + \sum_{p=2}^{n} \int_{p-1}^{p} \frac{dt}{t^{1+\epsilon}}) = \epsilon - (n^{-\epsilon} - 1) = 1 + \epsilon - n^{-\epsilon}$$

and

$$\sum_{p=1}^{n} \frac{\epsilon}{p^{1+\epsilon}} \quad \geq \quad \epsilon \sum_{p=1}^{n} \int_{p}^{p+1} \frac{dt}{t^{1+\epsilon}} = 1 - (n+1)^{-\epsilon}$$

We end the proof in this situation by noting that

$$\frac{1}{\epsilon}\ (1 + \epsilon - n^{-\epsilon}) = 1 + \frac{1 - n^{-\epsilon}}{\epsilon} = 1 + \frac{n^{\epsilon} - 1}{\epsilon n^{\epsilon}} = 1 + \frac{d_{\epsilon}(n)}{n^{\epsilon}}$$

and

$$\begin{aligned} \frac{1}{\epsilon}\ (1 - (n+1)^{-\epsilon}) \quad &= \quad \frac{d_{\epsilon}(n+1)}{(n+1)^{\epsilon}} \geq \frac{d_{\epsilon}(n)}{(n+1)^{\epsilon}} \\ &\geq \quad \frac{d_{\epsilon}(n)}{n^{\epsilon}} \left(\frac{n}{n+1}\right)^{\epsilon} \geq \frac{1}{2^{\epsilon}} \frac{d_{\epsilon}(n)}{n^{\epsilon}} \geq \frac{1}{2} \frac{d_{\epsilon}(n)}{n^{\epsilon}} \end{aligned}$$

In much the same way we have that for $0 < \epsilon \leq 1$,

$$\sum_{p=1}^{n} \frac{\epsilon}{p^{1-\epsilon}} \quad \leq \quad \epsilon\ (1 + \sum_{p=2}^{n} \int_{p-1}^{p} \frac{dt}{t^{1-\epsilon}}) = \epsilon + (n^{\epsilon} - 1)$$

and

$$\sum_{p=1}^{n} \frac{\epsilon}{p^{1-\epsilon}} \quad \geq \quad \epsilon \sum_{p=1}^{n} \int_{p}^{p+1} \frac{dt}{t^{1-\epsilon}} = (n+1)^{\epsilon} - 1$$

This implies that

$$d_{\epsilon}(n) \leq d_{\epsilon}(n+1) \leq \sum_{p=1}^{n} \frac{1}{p^{1-\epsilon}} \leq 1 + d_{\epsilon}(n)$$

On the other hand we note that

$$d_{\epsilon}(n) = \frac{1 - n^{-\epsilon}}{\epsilon n^{-\epsilon}} = \frac{n^{-\epsilon} - 1}{(-\epsilon) n^{-\epsilon}} = \frac{d_{-\epsilon}(n)}{n^{-\epsilon}}$$

Therefore

$$\frac{d_{-\epsilon}(n)}{n^{-\epsilon}} \leq \sum_{p=1}^{n} \frac{1}{p^{1-\epsilon}} \leq 1 + \frac{d_{-\epsilon}(n)}{n^{-\epsilon}}$$

and the proof of (18) is complete for any $\epsilon \in [-1, 1] - \{0\}$. For $\epsilon = 0$ we can take the limits $\epsilon \to 0$ in the above estimates or simply check directly that

$$\frac{1}{2} \log n \leq \log (n+1) \leq \sum_{p=1}^{n} \frac{1}{p} \leq 1 + \log n$$

This ends the proof of the lemma. ■

2 Asymptotic behavior

This section is mainly concerned with the proof of theorem 1.1 and theorem 1.2. We have chosen to separate the analysis in two parts. In a first subsection 2.1 we examine SIMC models with linear interactions. We begin with a model of ϵ-interacting random variables in which most of all calculations can be done explicitly. As we mentioned earlier in this example the decays to equilibrium presented in theorem 1.1 are sharp. The general linear SIMC model is treated in subsection 2.1.2. The proof of the estimates is essentially identical to the one of the ϵ-interacting sequence. Section 2.2 discusses the long time behavior of the two classes of non-linear SIMC models (1) and (6). Applications are presented in the further development of section 3.

2.1 Linear self-interacting models

We consider SIMC models (1) associated to a linear mapping of the form

$$\Phi(m) = mK$$

where K is a Markov transition on E. These models have a linear structure which makes it easy to analyze in details their asymptotic behavior in terms of the contraction coefficient $\beta(K)$ of the underlying transition K. Under the assumption that $\beta(K) < 1$ the limiting distribution is the unique invariant measure μ of the Markov transition K. If now we let

$$\mathbb{E}(S_n(f)) = \overline{S}_n(f)$$

denote the expected occupation measure associated to the SIMC model (1) with initial distribution η then we will check that

$$\|\overline{S}_n - \mu\|_{tv} \leq \frac{2}{n^{1-\beta(K)}} \; \|\eta - \mu\|_{tv}$$

Note that it is not plausible to obtain useful convergence results when $\beta(K) = 1$. For instance if $K(x, dy) = \delta_x(dy)$ then the whole chain is stuck in the initial point and $S_n = \delta_{X_0}$.

This section is organized as follows. In a first subsection we analyze the linear SIMC model associated to the Markov transition

$$K(x, dy) = \epsilon \; \delta_x(dy) + (1 - \epsilon) \; \mu(dy)$$

The resulting SIMC model simply consists in a sequence of ϵ-interacting random variables. The general linear SIMC model is treated in section 2.1.2. We also provide examples of transitions K such that $\beta(K) < 1$.

2.1.1 A sequence of ϵ-interacting variables

Let (μ, η) be a pair of probability measures on E and let $\epsilon \in [0, 1]$. In this section we consider the SIMC model X_n with initial distribution η and elementary transitions

$$\mathbb{P}(X_{n+1} \in dx \; |F_n \;) = \epsilon \; S_n(dx) + (1 - \epsilon) \; \mu(dx)$$

With a probability ϵ the particle X_{n+1} returns randomly and uniformly to one of the previous states X_p, $0 \leq p \leq n$ and with a probability $(1 - \epsilon)$ it chooses a new independent site according to the distribution μ. This model correspond to the situation (1) with

$$\Phi(m) = mK_\epsilon \quad \text{and} \quad K_\epsilon(x, dy) = \epsilon\ \delta_x(dy) + (1-\epsilon)\ \mu(dy)$$

We also notice that the measure μ is the unique fixed point of the mapping Φ. When the parameter ϵ is null and $\eta = \mu$ the chain reduces to a sequence of independent and identically distributed random variables with distribution μ. In this special case and by the law of large numbers the occupation measure S_n converges to μ as the time parameter tends to infinity. For instance we have that

$$\mathbb{E}((S_n(f) - \mu(f))^2) = \frac{1}{n+1}\ \sigma_\mu^2(f) \quad \text{with} \quad \sigma_\mu(f) = \mu((f-\mu(f))^2)^{1/2}$$

When the interaction parameter $\epsilon \in (0,1)$ it is natural to expect that the occupation measure S_n still converges to the fixed point μ but with a rate which depends on ϵ. It is convenient at this point to make a couple of remarks: First of all we notice that at each time the chain has a probability $(1-\epsilon)$ to visit a new random site according to μ and with a probability ϵ it chooses an occupied site with a probability proportional to the number of time the latter has been visited. In this sense the chain is attracted by sites which have been visited several times. We also notice that if $\varepsilon \leq 1/2$ then the chain has more chance to visit independent random sites with probability μ. In the opposite it tends to return randomly to the previous ones. In this sense ϵ measures the degree of interaction in the SIMC model.
The second remark is that the parameter ϵ also characterizes the contraction of the mapping Φ and the Dobrushin's coefficient of the Markov transition K_ϵ. More precisely for any $m_1, m_2 \in \mathcal{P}(E)$ we have that

$$\Phi(m_1) - \Phi(m_2) = \epsilon\ (m_1 - m_2) \quad \text{and} \quad \beta(K_\epsilon) = \epsilon$$

Proposition 2.1 *For any $n \geq 1$ we have*

$$\overline{S}_n - \mu \ = \ a(n)\ (\eta - \mu)$$

with

$$a(n) = \prod_{p=1}^{n} \frac{p+\epsilon}{p+1} \quad \textit{and} \quad 1/(2e) \leq n^{1-\epsilon}\ a(n) \leq 2 \tag{21}$$

Assume that $\eta = \mu$ and put $\delta = 2\epsilon - 1 (\in [-1,1])$. In this case there exists some finite constant $c \geq 1$ such that

$$\frac{1}{c}\ d_\delta(n) \leq n\ \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq c\ (d_\delta(n) + n^\delta) \tag{22}$$

for any measurable function f such that $\sigma_\mu(f) = 1$.

This proposition shows that there exits three types of behavior depending if ϵ is greater or lower or equal to $1/2$.

$$\begin{aligned}
\delta > 0 \ &\Rightarrow\ \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq \mathrm{c}/n^{1-\delta} \\
\delta < 0 \ &\Rightarrow\ \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq \mathrm{c}/(|\delta| n) \\
\delta = 0 \ &\Rightarrow\ \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq \mathrm{c}\ \log(n)/n
\end{aligned}$$

Proof: To prove (21) we first assume that $\mu(f) = 0$. In this case we have

$$\overline{S}_n(f) = \frac{n}{n+1}\overline{S}_{n-1}(f) + \frac{\epsilon}{n+1}\overline{S}_{n-1}(f) \;=\; \frac{n+\epsilon}{n+1}\,\overline{S}_{n-1}(f) \;=\; a(n)\,\eta(f)$$

Consequently for all bounded measurable functions f

$$\overline{S}_n(f) - \mu(f) = \overline{S}_n(f - \mu(f)) = a(n)\,(\eta(f) - \mu(f))$$

To prove (22) we introduce the decomposition

$$S_n(f) - \overline{S}_n(f) = \frac{n}{n+1}(S_{n-1}(f) - \overline{S}_{n-1}(f)) + \frac{1}{n+1}(f(X_n) - \overline{S}_{n-1}K_\epsilon(f))$$

By definition of K_ϵ we have for any $m_1, m_2 \in \mathcal{P}(E)$

$$(m_1K_\epsilon - m_2K_\epsilon) = \epsilon\,(m_1 - m_2)$$

Consequently

$$\begin{array}{l}\mathbb{E}\,(\,[S_{n-1}(f) - \overline{S}_{n-1}(f)]\,[f(X_n) - \overline{S}_{n-1}K_\epsilon(f)]\,)\\ = \mathbb{E}(\,[S_{n-1}(f) - \overline{S}_{n-1}(f)]\,[S_{n-1}K_\epsilon(f) - \overline{S}_{n-1}K_\epsilon(f)]\,) = \epsilon\,\mathbb{E}((S_{n-1}(f) - \overline{S}_{n-1}(f))^2)\end{array}$$

Therefore if we put

$$I_n(f) = (n+1)^2\,\mathbb{E}((S_n(f) - \overline{S}_n(f))^2)$$

then we have

$$I_n(f) = (1 + 2\epsilon/n)\,I_{n-1}(f) + \,\mathbb{E}(\,[f(X_n) - \overline{S}_{n-1}K_\epsilon(f)]^2)$$

Now we notice that

$$\mathbb{E}(\,[f(X_n) - \overline{S}_{n-1}K_\epsilon(f)]^2) = \overline{S}_{n-1}K_\epsilon(f^2) - (\overline{S}_{n-1}K_\epsilon(f))^2$$

Using (21) we find that

$$\begin{array}{rcl}\overline{S}_nK_\epsilon(f^2) & = & \epsilon\,\overline{S}_n(f^2) + (1-\epsilon)\,\mu(f^2)\\ & = & \epsilon\,a(n)\,\eta(f^2) + \epsilon\,(1 - a(n))\,\mu(f^2) + (1-\epsilon)\,\mu(f^2)\\ & = & \epsilon a(n)\,\eta(f^2) + (1 - \epsilon a(n))\,\mu(f^2)\end{array}$$

and

$$[\overline{S}_nK_\epsilon(f)]^2 = [\epsilon a(n)\,\eta(f) + (1 - \epsilon a(n))\,\mu(f)]^2$$

If we combine these two expressions we obtain

$$\mathbb{E}(\,[f(X_{n+1}) - \overline{S}_nK_\epsilon(f)]^2)$$

$$= \epsilon a(n)\sigma_\eta(f) + (1 - \epsilon a(n))\sigma_\mu(f) + \epsilon a(n)(1 - \epsilon a(n))\,(\eta(f) - \mu(f))^2$$

If we take $\eta = \mu$ we readily obtain

$$\mathbb{E}(\,[f(X_{n+1}) - \overline{S}_nK_\epsilon(f)]^2) = \sigma_\mu(f)$$

If we choose f such that $\mathrm{osc}(f) \leq 1$ we immediately obtain

$$I_n(f) \;\leq\; (1 + 2\epsilon/n)\;I_{n-1}(f) + 1 = \sum_{p=0}^{n}\;I_{p,n}(f)$$

with

$$I_{p,n}(f) = \prod_{q=p+1}^{n} (1 + 2\epsilon/q)$$

To end the proof of the proposition we use the following technical lemma 1.4. By (17) we have for any $n \geq 1$

$$\begin{aligned}
a(n) &= \prod_{p=1}^{n} \left(1 - \frac{(1-\epsilon)}{p+1}\right) = \prod_{p=2}^{n+1} \left(1 - \frac{(1-\epsilon)}{p}\right) \leq \left(\frac{2}{n+2}\right)^{1-\epsilon} \\
&\leq 2^{1-\epsilon} \left(\frac{n}{n+2}\right)^{1-\epsilon} \frac{1}{n^{1-\epsilon}} \leq \frac{2}{n^{1-\epsilon}}
\end{aligned}$$

and

$$\begin{aligned}
a(n) &= \prod_{p=2}^{n+1} \left(1 - \frac{(1-\epsilon)}{p}\right) \geq e^{-(1-\epsilon)} \left(\frac{1}{n+1}\right)^{1-\epsilon} \\
&\geq \frac{1}{e} \left(\frac{n}{n+1}\right)^{1-\epsilon} \frac{1}{n^{1-\epsilon}} \geq \frac{1}{2^{1-\epsilon} e} \frac{1}{n^{1-\epsilon}} \geq \frac{1}{2e} \frac{1}{n^{1-\epsilon}}
\end{aligned}$$

This ends the proof of (21). By (16) we have for any $1 \leq p \leq n$

$$e^{-4s} \; ((n+1)/(p+1))^{2\epsilon} \; \leq I_{p,n}(f) \leq (n/p)^{2\epsilon}$$

For $p = 0$ we use the fact that

$$\frac{e^{-4s}}{4} \; (n+1)^{2\epsilon} \leq \frac{e^{-4s}}{2^{2\epsilon}} \; (n+1)^{2\epsilon} \leq I_{0,n}(f) = I_{1,n}(f) \; (1+2\epsilon) \leq 3 \; n^{2\epsilon}$$

If we put $\delta = 2\epsilon - 1$ then by (18) we obtain

$$\begin{aligned}
I_n(f)/(n+1)^2 &\leq \frac{1}{n^{1-\delta}} \left(3 + \sum_{p=1}^{n} \frac{1}{p^{1+\delta}}\right) \\
&\leq \frac{1}{n^{1-\delta}} \left(4 + \frac{d_\delta(n)}{n^\delta}\right) = \frac{1}{n} \left(4n^\delta + d_\delta(n)\right)
\end{aligned}$$

and

$$\begin{aligned}
I_n(f)/(n+1)^2 &\geq \frac{e^{-4s}}{4} \frac{1}{(n+1)^{1-\delta}} \left(1 + \sum_{p=2}^{n+1} \frac{1}{p^{1+\delta}}\right) \\
&\geq \frac{e^{-4s}}{4} \frac{1}{n^{1-\delta}} \; \left(\frac{n}{n+1}\right)^{1-\delta} \; \sum_{p=1}^{n} \frac{1}{p^{1+\delta}} \\
&\geq \frac{e^{-4s}}{4 2^{1-\delta}} \frac{1}{n^{1-\delta}} \; \frac{1}{2} \; \frac{d_\delta(n)}{n^\delta} \geq \frac{e^{-4s}}{2^5} \; \frac{d_\delta(n)}{n}
\end{aligned}$$

■

2.1.2 General linear models

In this section we analyze the asymptotic behavior of an abstract SIMC model associated to a linear mapping Φ. We suppose this mapping is defined as follows

$$\Phi(m) = mK \tag{23}$$

for some Markov transition K on E such that $\beta(K) < 1$. Again we recall that this condition guarantees the existence and the uniqueness of a fixed point

$$\mu = \mu K \in \mathcal{P}(E)$$

Using the alternative representation (14) of the contraction coefficient $\beta(K)$ stated in lemma 1.3 it is easily verified that the linear mapping (23) satisfies condition **(HΦ)** with

$$\begin{array}{rcl} \mathcal{F} & = & \mathrm{Bosc}_1(E) \\ H(f,\mu) & = & \{\ [K(f) - \mu K(f)]/\beta(K)\ \} \quad \text{and} \quad \beta_{\mathcal{F}}(\Phi) = \beta(K) \end{array}$$

The contraction coefficient $\beta(K)$ is tied to mixing properties of the Markov transition K. For instance suppose there exists a probability measure $\nu \in \mathcal{P}(E)$ and a positive number $\epsilon \in (0,1]$ such that for any $(x,A) \in (E \times \mathcal{E})$

$$K(x,A) \ \geq \ (1-\epsilon)\ \nu(A) \tag{24}$$

Then, by (13) we have that $\beta(K) \leq \epsilon$. Next we present two examples which indicate that this condition is more connected to mixing properties of K rather than compactness properties of the state space.

Example 1 *Let $E = \mathbb{R}$ and let $a : \mathbb{R} \to \mathbb{R}$ be a bounded Borel function. One can check that the bi-Laplace transition*

$$K(x,dy) = \frac{c}{2}\ e^{-c|y-a(x)|}\ dy$$

with $0 < c < \infty$ satisfies (24) with

$$\nu(dy) = \frac{c}{2}\ e^{-c|y|}\ dy \quad \text{ and } \quad \epsilon = 1 - e^{-c\|a\|_\infty}$$

Example 2 *Let $E = \{x_i\ ;\ i \in I\}$ be a collection of states indexed by a countable set I. Let K be the Markov transition defined by*

$$K(x,dy) = \sum_{i\in I}\ a_i(x)\ \delta_{x_i}$$

for some positive functions $\{a_i\ ;\ i \in I\}$ on the set E such that $\sum_{i\in I}\ a_i = 1$. Recalling that

$$\|K(x,\centerdot) - K(y,\centerdot)\|_{tv} = \frac{1}{2}\sum_{i\in I}|K(x,x_i) - K(y,x_i)|$$

by (12) we find that $\beta(K) < 1$ as soon as $\sum_{i\in I}\ \mathrm{osc}(a_i) < 2$.

Proposition 2.2 *For any $n \geq 1$ we have the estimate*

$$\|\overline{S}_n - \mu\|_{tv} \leq \frac{2}{n^{1-\beta(K)}} \ \|\eta - \mu\|$$

In addition for any $f \in Osc_1(E)$ we have

$$n \ \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq c \ (d_{\delta(K)}(n) + n^{\delta(K)}) \quad \textit{with} \quad \delta(K) = 2\beta(K) - 1$$

As observed in section 2.1.1 the above estimates induce three different types of decays

$$\begin{array}{lll} \beta(K) > 1/2 & \Rightarrow & \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq \mathrm{c}/n^{2(1-\beta(K))} \\ \beta(K) < 1/2 & \Rightarrow & \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq \mathrm{c}/(|1 - 2\beta(K)|n) \\ \beta(K) = 1/2 & \Rightarrow & \mathbb{E}((S_n(f) - \overline{S}_n(f))^2) \leq \mathrm{c} \ \log(n)/n \end{array}$$

From proposition 2.1 we recall that these estimates cannot be improved without adding some some additional regularity condition on the Markov transition K.
Proof: We use the fact that $\overline{S}_n$ can be represented in the form

$$\overline{S}_n = \frac{n}{n+1}\overline{S}_{n-1} + \frac{1}{n+1}\overline{S}_{n-1}K = \overline{S}_{n-1}K_n$$

with

$$K_n(x, dy) = \frac{n}{n+1} \ \delta_x(dy) + \frac{1}{n+1} \ K(x, dy)$$

It is also easy to check that $\mu = \mu K_n$. It follows that

$$\overline{S}_n = \eta K_1 \ldots K_n \quad \text{and} \quad \mu = \mu K_1 \ldots K_n$$

By definition of Dobrushin's contraction coefficient we see that

$$\|\overline{S}_n - \mu\|_{tv} \leq \prod_{p=1}^{n} \beta(K_p) \ \|\eta - \mu\|_{tv}$$

Since we have for each $n \geq 1$

$$\beta(K_n) \leq \frac{n + \beta(K)}{n+1}$$

we conclude that

$$\|\overline{S}_n - \mu\|_{tv} \leq \prod_{p=1}^{n} \frac{p + \beta(K)}{p+1} \ \|\eta - \mu\|_{tv}$$

The first statement is now easily proved by applying lemma 1.4. More precisely using (17) we have that

$$\prod_{p=1}^{n} \frac{p + \beta(K)}{p+1} = \prod_{p=2}^{n+1} \left(1 - \frac{1 - \beta(K)}{p}\right) \leq \left(\frac{2}{n+2}\right)^{1-\beta(K)} \leq \frac{2}{n^{1-\beta(K)}}$$

To prove the second assertion we choose $f \in \mathrm{Osc}_1(E)$ and we note

$$I_n(f) = (n+1)^2 \ \mathbb{E}((S_n(f) - \overline{S}_n(f))^2)$$

By definition of the SIMC we have

$$I_n(f) = \ I_{n-1}(f) + \ \mathbb{E}(\ [f(X_n) - \overline{S}_{n-1}K_\epsilon(f)]^2) + 2n \ J_{n-1}(f, Kf) \tag{25}$$

with

$$J_n(f,g) = \mathbb{E}([S_n(f) - \overline{S}_n(f)][S_n(g) - \overline{S}_n(g)])$$

We observe that for any $x, y \in E$ and $m \in \mathcal{P}(E)$

$$|f(x) - m(f)| \leq \int |f(x) - f(y)| \ m(dy) \leq 1$$

and

$$\begin{aligned} |K(f)(x) - K(f)(y)| &\leq \sup\{|K(g)(x) - K(g)(y)| \ ; \ g \in \mathrm{Osc}_1(E)\} \\ &= \|K(x,.) - K(y,.)\|_{tv} \leq \beta(K) \end{aligned}$$

From these observations it follows that

$$\mathbb{E}(\ [f(X_n) - \overline{S}_{n-1}K_\epsilon(f)]^2) \leq 1$$

and

$$f \in \mathrm{Osc}_1(E) \Longrightarrow \frac{1}{\beta(K)} \ K(f) \ \in \mathrm{Osc}_1(E)$$

On the other hand, by Cauchy-Schwartz's inequality

$$(n+1)^2 \ J_n(f, Kf) \leq \beta(K) \ I_n(f)^{1/2} \ I_n(Kf/\beta(K))^{1/2}$$

Therefore if we write

$$I_n = \sup\{I_n(f) \ ; \ f \in \mathrm{Osc}_1(E)\}$$

we conclude from (25) that

$$I_n \leq (1 + 2\beta(K)/n) \ I_{n-1} + 1$$

We complete the proof using the same arguments as the end of the proof of proposition 2.1. ■

2.2 Non-linear self-interacting models

In this section we analyze the long time behavior of the SIMC models (1) and (6) associated respectively to an abstract mapping $\Phi : \mathcal{P}(E) \to \mathcal{P}(E)$ and to a collection of Markov transition $\{K_m \ ; \ m \in \mathcal{P}(E)\}$.
First we note that taking suprema over $\mathcal{F}$ in **(HΦ)**, we obtain that

$$\forall \ \mu, \nu \in \mathcal{P}(E), \qquad d_{\mathcal{F}}(\Phi(\mu), \Phi(\nu)) \ \leq \ \beta_{\mathcal{F}}(\Phi) d(\mu, \nu)$$

so that an usual fixed point theorem shows there exists a unique $\mu \in \mathcal{P}(E)$ verifying $\mu = \Phi(\mu)$.

2.2.1 Proof of theorem 1.1

The aim of this section is to prove theorem (1.1). We suppose condition **(HΦ)** is met for some class of functions $\mathcal{F}$ and we denote by S_n the occupation measure at time n associated to the SIMC model (1). For each $f \in \mathcal{F}$ we write

$$I_n(f) = (n+1)^2 \ \mathbb{E}((S_n(f) - \mu(f))^2) \quad \text{and} \quad I_n = \sup\{I_n(f)\ ;\ f \in \mathcal{F}\}$$

where $\mu = \Phi(\mu)$ stands for the unique fixed point of the mapping Φ. Using the decomposition

$$S_n(f) - \mu(f) = \frac{n}{n+1}\ [S_{n-1}(f) - \mu(f)] + \frac{1}{n+1}\ [f(X_n) - \Phi(\mu)(f)]$$

we first show that

$$I_n(f) = \ I_{n-1}(f) + \ \mathbb{E}(\ [f(X_n) - \Phi(\mu)(f)]^2) + 2n\ C_{n-1}(f) \tag{26}$$

with

$$C_n(f) = \mathbb{E}([S_n(f) - \mu(f)]\ [\Phi(S_n)(f) - \Phi(\mu)(f)])$$

Under our assumptions we have that

$$|\Phi(S_n)(f) - \Phi(\mu)(f)| \leq \sum_{h \in H(f,\mu)} b_{f,\mu}(h)\ |S_n(h) - \mu(h)|$$

This implies that

$$|C_n(f)| \ \leq \ \sum_{h \in H(f,\mu)} b_{f,\mu}(h)\ \mathbb{E}(|S_n(f) - \mu(f)|\ |S_n(h) - \mu(h)|)$$

By Cauchy-Schwartz's inequality and taking the supremum over all $f \in \mathcal{F}$ in the right hand side we obtain

$$(n+1)\ |C_n(f)| \ \leq \ \beta_{\mathcal{F}}(\Phi)\ I_n$$

Finally by (26) we conclude that

$$\begin{aligned} I_n \ &\leq \ (1 + 2\beta_{\mathcal{F}}(\Phi)/n)\ I_{n-1}(f) + 2 \\ &\leq \ \prod_{p=1}^{n} \left(1 + \frac{2\beta_{\mathcal{F}}(\Phi)}{p}\right) \eta[(f - \mu(f))^2] + 2\sum_{p=1}^{n} \prod_{q=p+1}^{n} \left(1 + \frac{2\beta_{\mathcal{F}}(\Phi)}{q}\right) \end{aligned}$$

The end of the proof now follows the same lines of arguments as the end of proof of proposition 2.1. ■

2.2.2 Proof of theorem 1.2

In this section we suppose condition **(HK)** is met for some class of functions $\mathcal{F}$ and we denote by S_n the occupation measure at time n associated to the SIMC model (6). Under **(HK)** there exists for each m a K_m-invariant measure

$$\pi(m) = \pi(m)K_m \in \mathcal{P}(E)$$

We associate to the mapping $\pi : \mathcal{P}(E) \to \mathcal{P}(E)$ and to the distribution flow S_n the random measures

$$S_n^{\pi} = \frac{1}{n+1} \sum_{q=0}^{n} \ \pi(S_q)$$

Furthermore, in quite the same way as in the beginning of subsection 2.2, we see there exists a unique fixed point $\mu = \pi(\mu)$.

Proposition 2.3 *For each $f \in \mathcal{F}$ and $p > 1$ we have that*

$$\sup_{n \geq 1} \sqrt{n}\ \mathbb{E}(|S_n(f) - S_n^\pi(f)|^p)^{1/p} < \infty$$

Proof: Given a function $f \in \mathcal{F}$ and a distribution m on E we have a solution g_m of the Poisson's equation

$$f - \pi(m)(f) = g_m - K_m(g_m)$$

and it is given by

$$g_m = \sum_{n \geq 0} K_m^n(f - \pi(m)(f))$$

Furthermore under our assumptions we have that

$$\|g_m\|_\infty \leq 2 \sum_{n \geq 0} \epsilon(n)$$

For any $n \geq 0$ we notice that

$$\sum_{p=0}^{n} [f(X_p) - \pi(S_p)(f)] = \sum_{p=0}^{n} [g_{S_p}(X_p) - K_{S_p}(g_{S_p})(X_p)]$$

Using the decomposition

$$\begin{aligned} g_{S_p}(X_p) - K_{S_p}(g_{S_p})(X_p) &= [g_{S_p}(X_p) - g_{S_{p+1}}(X_{p+1})] \\ &\quad + [g_{S_{p+1}}(X_{p+1}) - g_{S_p}(X_{p+1})] \\ &\quad + [g_{S_p}(X_{p+1}) - K_{S_p}(g_{S_p})(X_p)] \end{aligned}$$

we obtain

$$\begin{aligned} &\textstyle\sum_{p=0}^{n} [f(X_p) - \pi(S_p)(f)] \\ &= -(g_{S_{n+1}}(X_{n+1}) - g_{S_0}(X_0)) + \textstyle\sum_{p=0}^{n} (g_{S_{p+1}} - g_{S_p})(X_{p+1}) + M_n(f) \end{aligned}$$

with the F-martingale

$$M_n(f) = \sum_{p=0}^{n} [g_{S_p}(X_{p+1}) - K_{S_p}(g_{S_p})(X_p)]$$

Note that the quadratic variation of the latter is given by

$$[M(f)]_n = \sum_{p=0}^{n} [g_{S_p}(X_{p+1}) - K_{S_p}(g_{S_p})(X_p)]^2 \leq c\ (n+1)$$

Using Burkholder's inequality (cf. for instance (27), p.499 in [20]) we have for any $p > 1$

$$\mathbb{E}(|M_n(f)|^p)^{1/p} \leq c(p)\ \sqrt{n+1}$$

for some finite constant $c(p) < \infty$ which depends on the parameter p. In view of (9) and (10) we have for any pair (m_1, m_2) of distributions

$$\begin{aligned} \|g_{m_1} - g_{m_2}\|_\infty &\leq \sum_{n \geq 0} \epsilon(n)\ [|\pi(m_1)(f) - \pi(m_2)(f)| + \|m_1 - m_2\|_{tv}] \\ &\leq c\ \|m_1 - m_2\|_{tv} \end{aligned}$$

with $c = (1 + 2\beta_{\mathcal{F}}(\pi)) \sum_{n\geq 0} \epsilon(n) < \infty$. Since $\|S_{p+1} - S_p\| \leq 2/(p+2)$ we conclude that

$$\|g_{S_{p+1}} - g_{S_p}\|_\infty \leq 2c/(p+2)$$

and therefore

$$|\sum_{p=0}^{n} (g_{S_{p+1}} - g_{S_p})(X_{p+1})| \leq c \ \log(n+1)$$

If we combine the above estimates we obtain for sufficiently large values of the time parameter

$$\mathbb{E}(|\sum_{p=0}^{n} [f(X_p) - \pi(S_p)(f)]|^p)^{1/p} \leq c'(p) \ \sqrt{n+1}$$

for some finite constant $c'(p) < \infty$ which depends on the parameter p. This ends the proof of the proposition. ■

Under our assumptions the mapping $\pi : \mathcal{P}(E) \to \mathcal{P}(E)$ has a unique fixed point

$$\mu = \pi(\mu) \in \mathcal{P}(E)$$

In the next proposition we estimate the convergence decays of the flow S_n^π towards this equilibrium measure.

Proposition 2.4 *For any $f \in \mathcal{F}$ we have three different types of $\mathbb{L}_2$-mean error decays:*

$$\begin{array}{lcl}
\beta_{\mathcal{F}}(\pi) > 3/4 & \Rightarrow & n^{2(1-\beta_{\mathcal{F}}(\pi))} \ \mathbb{E}((S_n^\pi(f) - \mu(f))^2) \leq c \\
\beta_{\mathcal{F}}(\pi) < 3/4 & \Rightarrow & \sqrt{n} \ \mathbb{E}((S_n^\pi(f) - \mu(f))^2) \leq c \\
\beta_{\mathcal{F}}(\pi) = 3/4 & \Rightarrow & \dfrac{\sqrt{n}}{\log(n)} \mathbb{E}((S_n^\pi(f) - \mu(f))^2) \leq c
\end{array}$$

Proof: Let $f \in \mathcal{F}$ be given. For each $n \geq 0$ we define

$$I_n(f) = S_n^\pi(f) - \mu(f) = \frac{1}{n+1} \sum_{p=0}^{n} [\pi(S_p)(f) - \pi(\mu)(f)] \quad \text{and} \quad J_n = \sup_{f\in\mathcal{F}} \mathbb{E}(I_n(f)^2)$$

We also introduce the decomposition

$$I_n(f) \ = \ \frac{n}{n+1} \ I_{n-1}(f) \ + \ \frac{1}{n+1} \ [\pi(S_n)(f) - \pi(\mu)(f)] \tag{27}$$

Using (9) we have

$$|\pi(S_n)(f) - \pi(\mu)(f)| \leq \sum_{h\in H(f,\mu)} b_{f,\mu}(h) \ |S_n(h) - \mu(h)| \tag{28}$$

for some subset $H(f,\mu) \subset \mathcal{F}$ and some collection of numbers $b_{f,\mu}(h) \in [0,1)$, $h \in H(f,\mu)$ such that

$$\beta_{\mathcal{F}}(\pi) = \sup\{ \sum_{h\in H(f,\mu)} b_{f,\mu}(h) \ ; \ f \in \mathcal{F}\} \in (0,1)$$

Using the decompositions

$$\begin{aligned} S_n - \mu &= [S_n - S_n^\pi] + [S_n^\pi - \mu] \\ S_n^\pi - \mu &= \frac{n}{n+1}[S_{n-1}^\pi - \mu] + \frac{1}{n+1}[\pi(S_n) - \pi(\mu)] \\ &= \frac{n}{n+1} \ I_{n-1} \ + \ \frac{1}{n+1} \ [\pi(S_n) - \pi(\mu)] \end{aligned}$$

and from the inequality (28) we obtain

$$\begin{aligned} |\pi(S_n)(f) - \pi(\mu)(f)| \ &\leq \sum_{h \in H(f,\mu)} b_{f,\mu}(h) \ |S_n(h) - S_n^\pi(h)| \\ &\quad + \frac{n}{n+1} \sum_{h \in H(f,\mu)} b_{f,\mu}(h) \ |I_{n-1}(h)| \\ &\quad\quad + \frac{1}{n+1} \sum_{h \in H(f,\mu)} b_{f,\mu}(h) \ |\pi(S_n)(h) - \pi(\mu)(h)| \\ &\leq \frac{c}{n+1} + \sum_{h \in H(f,\mu)} b_{f,\mu}(h) \ [|S_n(h) - S_n^\pi(h)| + |I_{n-1}(h)|] \end{aligned}$$

Using proposition 2.3 and Cauchy-Schwartz's inequality we note that for sufficiently large n

$$\mathbb{E}(|I_{n-1}(f)| \ |S_n(h) - S_n^\pi(h)|) \leq c/\sqrt{n+1}$$

In much the same way we also check that for each $f, h \in \mathcal{F}$

$$\mathbb{E}(|I_{n-1}(f)| \ |I_{n-1}(h)|) \leq \mathbb{E}(|I_{n-1}(f)|^2)^{1/2} \mathbb{E}(|I_{n-1}(h)|^2)^{1/2}$$

and

$$\sum_{h \in H(f,\mu)} b_{f,\mu}(h) \ \mathbb{E}(|I_{n-1}(f)| \ |I_{n-1}(h)|) \ \leq \ \beta_{\mathcal{F}}(\pi) \ J_{n-1}$$

Using the decomposition (27) it is now easy to show that

$$J_n \leq (\frac{n}{n+1})^2 \ (1 + \frac{2\beta_{\mathcal{F}}(\pi)}{n}) \ J_{n-1} + \frac{c}{(n+1)^2} \ \sqrt{n+1}$$

Thus we have

$$J_n \ \leq \ \frac{c}{(n+1)^2} \sum_{p=0}^{n} \ [\prod_{q=p+1}^{n} (1 + \frac{2\beta_{\mathcal{F}}(\pi)}{q})] \ \sqrt{p+1}$$

Finally we use lemma 1.4 to demonstrate that

$$J_n \ \leq \ \frac{c}{n^{2(1-\beta_{\mathcal{F}}(\pi))}} \ (1 + \sum_{p=1}^{n} p^{\frac{1}{2} - 2\beta_{\mathcal{F}}(\pi)})$$

When $\beta_{\mathcal{F}}(\pi) \in (3/4, 1]$ we have $2\beta_{\mathcal{F}}(\pi) - \frac{1}{2} > 1$ and therefore $n^{2(1-\beta_{\mathcal{F}}(\pi))} \ J_n \leq c$. If $\beta_{\mathcal{F}}(\pi) \in [0, 3/4)$ then we have $2\beta_{\mathcal{F}}(\pi) - \frac{1}{2} < 1$ and

$$\sum_{p=1}^{n} p^{\frac{1}{2} - 2\beta_{\mathcal{F}}(\pi)} \ \leq \ c \ n^{1 - (2\beta_{\mathcal{F}}(\pi) - \frac{1}{2})}$$

In this situation we conclude that $\sqrt{n} \ J_n \leq c$. Finally if $\beta_{\mathcal{F}}(\pi) = 3/4$ then we have $2\beta_{\mathcal{F}}(\pi) - \frac{1}{2} = 1$ and $2(1 - \beta_{\mathcal{F}}(\pi)) = 1/2$. This yields that $\sqrt{n} \ J_n \leq c \ \log n$ and the proof of the proposition is now completed.

■

3 Applications

In this section we illustrate the verification of the regularity conditions **(HΦ)** and **(HK)** for some specific examples.

In a first subsection 3.1 we consider the SIMC model (1) associated to a mapping $\Phi : \mathcal{P}(E) \to \mathcal{P}(E)$. We first examine the genetic SIMC model (2). We propose a sufficient condition on the fitness function and the mutation kernel underwhich condition **(HΦ)** is met. We also illustrate this regularity condition for Gaussian and Poisson SIMC models.

In section 3.2 we consider the SIMC model (6) associated to a collection of Markov transitions $\{K_m \ ; \ m \in \mathcal{P}(E)\}$. Condition **(H$K$)** guarantees for each m the existence of a K_m-invariant measure $\pi(m) = \pi(m)K_m \in \mathcal{P}(E)$. We examine situations in which the mapping $\pi : \mathcal{P}(E) \to \mathcal{P}(E)$ is respectively constant, linear and non-linear. In the final subsection 3.2.3 we analyze the genetic SIMC model (8) with random selection and mutation transitions. We propose a sufficient condition on the triplet (a, g, K) underwhich condition **(HK)** is met. We also initiate a comparison between this genetic model and the genetic model (2).

3.1 Φ-interacting Markov chains

3.1.1 Genetic self-interactions

Consider the genetic SIMC model (3) associated to the mapping

$$\Phi(m)(f) = \frac{m(g\ (Kf))}{m(g)}$$

which we presented in the introductive section. Then we have the following proposition.

Proposition 3.1 *Suppose the potential function g and the mutation transition K are chosen so that*

$$g_{min} = 1 \qquad \textit{and} \qquad \sup_{x \in E} g(x) < \frac{1}{2}\left(\frac{1}{\beta(K)} + 1\right)$$

Then the mapping Φ satisfies condition **(HΦ)** *with*

$$\mathcal{F} = Bosc_1(E) \qquad \textit{and} \qquad \beta_{\mathcal{F}}(\Phi) = (1 + 2\text{osc}(g))\ \beta(K)$$

Remark 3.2: The first condition on the potential function is not really restrictive. When the infimum $g_{min} > 0$ we can replace in (2) the fitness function g by g/g_{min}. More precisely let $\Phi^{(h)}$ denotes the mapping from $\mathcal{P}(E)$ into itself defined as in (2) by replacing the function g by h. We notice that

$$\begin{aligned}(1 + 2\text{osc}(h/h_{min})) &= (2\text{osc}(h) + h_{min})/h_{min} \\ &= (\text{osc}(h) + \|h\|_\infty)/h_{min}\end{aligned}$$

From this observation we find that the mapping $\Phi^{(h)}$ satisfies **(HΦ)** with

$$\beta_{\mathcal{F}}(\Phi^{(h)}) = (\text{osc}(h) + \|h\|_\infty)\beta(K)/h_{min} \tag{29}$$

In addition we have

$$\sup_{x,y}\frac{h(x)}{h(y)} < \frac{1}{2}\ \left(\frac{1}{\beta(K)}+1\right) \Longrightarrow \beta_{\mathcal{F}}(\Phi^{(h)}) < 1$$

The above condition relates the oscillations of the potential function h with the mixing properties of the Markov mutation transition. The more K is mixing the more the potential is allowed to oscillate. Note that if $K(x,dy) = \delta_x(dy)$ then $\beta(K) = 1$ and this condition is never met. In fact in this degenerate situation the SIMC reduces to $X_n = X_0$ and the occupation measure $S_n = \delta_{X_0}$ which just trivially converges to δ_{X_0}!

Proof of proposition 3.1: To prove that the mapping Φ satisfies **(HΦ)** we use the decomposition

$$\Phi(\eta)(f) - \Phi(\mu)(f) = \frac{1}{\eta(g)}\ \left(\eta(g\ [K(f) - \Phi(\mu)(f)]) - \mu(g\ [K(f) - \Phi(\mu)(f)])\right)$$

Then we notice that

$$K(f)(x) - \Phi(\mu)(f) = \int\ (Kf(x) - Kf(y))\ \frac{g(y)}{\mu(g)}\ \mu(dy)$$

and

$$\operatorname{osc}(K(f))\ \leq\ \beta(K)\ \operatorname{osc}(f)$$

According to this observation if we set $f_\mu(x) = f(x) - \Phi(\mu)(f)$ then we find that

$$\begin{aligned}
|g(x)K(f_\mu)(x) - g(y)K(f_\mu)(y)| \ &\leq\ |g(x)-g(y)|\ |K(f_\mu)(x)| \\
&\qquad\qquad +g(y)\ |K(f)(x) - K(f)(y)| \\
&\leq\ \beta(K)\ (\ \operatorname{osc}(g) + \|g\|_\infty\ )\ \operatorname{osc}(f)
\end{aligned}$$

Since $\|g\|_\infty = \operatorname{osc}(g)+1$ one gets for any $f\in\mathcal{F}$

$$\operatorname{osc}(gK(f_\mu)) \leq \beta_{\mathcal{F}}(\Phi) \quad \text{and} \quad \|gK(f_\mu)\|\ \leq\ (\operatorname{osc}(g)+1)\ \beta(K) \leq \beta_{\mathcal{F}}(\Phi)$$

We conclude that condition **(HΦ)** is met with

$$H(f,\mu) = \{\ g\ [K(f) - \Phi(\mu)(f)]/\beta_{\mathcal{F}}(\Phi)\ \}$$

This end the proof of the proposition. ∎

3.1.2 Gaussian and Poisson self-interactions

In this short section we illustrate condition **(HΦ)** in the context of Gaussian and Poisson mean field interactions. Before presenting these two examples let us start with a generic and abstract situation. Suppose (E,d) is a metric space and let $(E',\mathcal{E}')$ be an auxiliary measurable space and $\gamma\in\mathcal{P}(E')$. Also suppose $\theta : E'\times\mathcal{P}(E)\to E$ is a measurable mapping satisfying the following condition.

$$|\theta(x',m_1) - \theta(x',m_2)| \leq \sum_{h\in H}\ b(h)\ |m_1(h) - m_2(h)|$$

for some subset $H \subset \mathrm{Lip}_1(E)$ and some collection of numbers $\{b(h)\ ;\ h \in H\}$ with $\beta = \sum_{h\in H} b(h) < 1$. We now associate to the pair (θ,γ) the non-linear mapping Φ defined as

$$\Phi(m)(f) = \int_{E'} f(\theta(x',m))\ \gamma(dx')$$

By the regularity condition on the mapping θ for any $f \in \mathrm{Lip}_1(E)$ we have that

$$|\Phi(m_1)(f) - \Phi(m_2)(f)| \leq \sum_{h\in H} b(h)\ |m_1(h) - m_2(h)|$$

Therefore condition **(HΦ)** holds with

$$\mathcal{F} = \mathrm{Lip}_1(E)\ , \qquad H(f,\mu) = H \quad \text{and} \quad \beta_{\mathcal{F}}(\Phi) = \beta$$

Gaussian interactions

Let $E = \mathbb{R}$ and let Φ be the Gaussian type non-linear mapping

$$\Phi(m)(dx) = \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}(x-m(a))^2}\ dx$$

for some $a \in \mathrm{Lip}_1(E)$ with $\|a\|_{\infty,\mathrm{L}} < 1$. Arguing as above we check that condition **(HΦ)** is satisfied with

$$\mathcal{F} = \mathrm{Lip}_1(E) \qquad H(f,\mu) = \{a/\|a\|_{\infty,\mathrm{L}}\} \quad \text{and} \quad \beta_{\mathcal{F}}(\Phi) = \|a\|_{\infty,\mathrm{L}} \tag{30}$$

Poisson interactions

Let $E = \mathbb{R}$ and let Φ be the Poisson type non-linear mapping

$$\Phi(m)(dx) = 1_{\mathbb{R}_+(x)}\ \frac{1}{m(a)}\ e^{-x/m(a)}\ dx$$

for some positive intensity function $a \in \mathrm{Lip}_1(E)$ with $\|a\|_{\infty,\mathrm{L}} < 1$. For any $f \in \mathrm{Lip}_1(E)$ we have that

$$|\phi(m_1)(f) - \phi(m_2)(f)| \leq \int u\ |m_1(a) - m_2(a)|\ e^{-u}\ du = |m_1(a) - m_2(a)|$$

Here again it is easily check that condition **(HΦ)** is satisfied with the same class of functions and parameters (30).

3.2 K_m-interacting Markov chains

3.2.1 Constant π-mapping

Suppose $K(x,dy)$ is a Markov transition on E with $\beta(K) < 1$ and let a, b be two measurable mappings from E into $[0,1]$ such that $a + b = 1$. We associate to the pair (a,b) the collection of Markov transitions

$$K_m(x,dy)\ =\ m(a)\ \delta_x(dy)\ +\ m(b)\ K(x,dy) \tag{31}$$

Proposition 3.3 *Suppose the triplet (a, b, K) satisfies the following condition*

$$\|a\|_\infty + \|b\|_\infty \ \beta(K) < 1$$

Then the Markov transition K has a unique invariant measure $\mu \in \mathcal{P}(E)$ and for each $m \in \mathcal{P}(E)$ and $n \geq 0$ we have

$$\pi(m) = \mu \quad \text{and} \quad K_m^n = \sum_{p=0}^{n} C_n^p \ m(a)^{n-p} \ m(b)^p \ K^p \tag{32}$$

In addition K_m satisfies $(\mathbf{H}K)$ *for the class of functions*

$$\mathcal{F} = \{f \in \mathcal{B}_b(E) \ ; \ \|f\|_\infty \leq 1\} \qquad \text{and} \qquad \beta_{\mathcal{F}}(\pi) = 0$$

Remark 3.4: Let S_n be the occupation measure associated to the SIMC model (6) with transitions (31). This proposition in conjunction with proposition 2.3 tells us that for each $f \in \mathcal{B}_b(E)$ and $p > 1$

$$\sup_{n \geq 1} \sqrt{n} \ \mathbb{E}(|S_n(f) - \mu(f)|^p)^{1/p} < \infty$$

Proof of proposition 3.3: Under our assumptions we observe that $\beta(K) < 1$. Thus, there exists a unique invariant measure $\mu = \mu K$ and for each $m \in \mathcal{P}(E)$ we clearly have that $\pi(m) = \mu$. From these observations condition (9) is trivially satisfied with for any class of functions $\mathcal{F}$ with $\beta_{\mathcal{F}}(\pi) = 0$. The expression (32) is proved by a simple induction on the time parameter n. Using the decomposition

$$K_m(f)(x) - K_m(f)(y) = m(a) \ (f(x) - f(y)) + m(b) \ (K(f)(x) - K(f)(y))$$

we see that

$$\text{osc}(K_m(f)) \leq [m(a) + m(b)\beta(K)] \ \text{osc}(f) \leq \beta_{a,b}(K) \ \text{osc}(f)$$

where

$$\beta_{a,b}(K) = \|a\|_\infty + \|b\|_\infty \ \beta(K)$$

This yields that

$$\sup_{m \in \mathcal{P}(E)} \beta(K_m) \leq \beta_{a,b}(K)$$

From (32) we have immediately for any pair of measures $(m_1, m_2) \in \mathcal{P}(E) \times \mathcal{P}(E)$ and for any function f with $\|f\|_\infty \leq 1$,

$$\begin{aligned} & \|K_{m_1}^n(f) - K_{m_2}^n(f)\|_\infty \\ & \leq \ 2 \ \textstyle\sum_{p+q=n} C_n^q \ \|a\|_\infty^p \ \|b\|_\infty^q \ \beta(K)^q \ |m_1(\overline{a})^p \ m_1(\overline{b})^q - m_2(\overline{a})^p \ m_2(\overline{b})^q| \end{aligned} \tag{33}$$

where

$$\overline{a} = a/\|a\|_\infty \qquad \text{and} \qquad \overline{b} = b/\|b\|_\infty$$

It is now straightforward to check that

$$\begin{aligned} |m_1(\overline{a})^p - m_2(\overline{a})^p| \quad &= \quad |m_1(\overline{a}) - m_2(\overline{a})| \ |\sum_{k=0}^{p-1} m_1(\overline{a})^k \ m_2(\overline{a})^{(p-1)-k}| \\ &\leq \quad 2p \ \|m_1 - m_2\|_{tv} \end{aligned}$$

and similarly

$$|m_1(\overline{b})^q - m_2(\overline{b})^q| \leq 2q \; \|m_1 - m_2\|_{tv}$$

Using the decomposition

$$\begin{aligned}
& m_1(\overline{a})^p \; m_1(\overline{b})^q - m_2(\overline{a})^p \; m_2(\overline{b})^q \\
& = (m_1(\overline{a})^p \; - m_2(\overline{a})^p) \; m_1(\overline{b})^q + m_2(\overline{a})^p \; (m_1(\overline{b})^q - m_2(\overline{b})^q)
\end{aligned}$$

the above estimates give

$$|m_1(\overline{a})^p \; m_1(\overline{b})^q - m_2(\overline{a})^p \; m_2(\overline{b})^q| \leq 2n \; \|m_1 - m_2\|_{tv}$$

for each $p + q = n$. From (33) this yields that

$$\|K^n_{m_1}(f) - K^n_{m_2}(f)\|_\infty \;\leq\; 4n \; \beta_{a,b}(K)^n \; \|m_1 - m_2\|_{tv}$$

Finally we notice that for any bounded measurable functions f_1, f_2 such that $\mu(f_1) = 0 = \mu(f_2)$, and for any $m \in \mathcal{P}(E)$

$$\begin{aligned}
\|K^n_m(f_1) - K^n_m(f_2)\|_\infty &= \|K^n_m([f_1 - f_2] - \mu([f_1 - f_2]))\|_\infty \\
&\leq \operatorname{osc}(K^n_m(f_1 - f_2)) \leq \beta(K^n_m) \operatorname{osc}(f_1 - f_2) \\
&\leq 2\beta_{a,b}(K)^n \; \|f_1 - f_2\|_\infty
\end{aligned}$$

This implies that

$$\begin{aligned}
\|K^n_{m_1}(f_1) - K^n_{m_2}(f_2)\|_\infty &\leq \|K^n_{m_1}(f_1) - K^n_{m_2}(f_1)\|_\infty + \|K^n_{m_2}(f_1 - f_2)\|_\infty \\
&\leq 4n \; \beta_{a,b}(K)^n \; \|m_1 - m_2\|_{tv} + 2\beta_{a,b}(K)^n \; \|f_1 - f_2\|_\infty \\
&\leq 4n \; \beta_{a,b}(K)^n \; [\|m_1 - m_2\|_{tv} + \|f_1 - f_2\|_\infty]
\end{aligned}$$

for any f_1, f_2 such that $\mu(f_1) = 0 = \mu(f_2)$ and $\|f_1\|_\infty, \|f_2\|_\infty \leq 1$. We conclude that (10) is satisfied with $\epsilon(n) = 4n \; \beta_{a,b}(K)^n$. ∎

3.2.2 Non-linear π-mapping

Consider next an abstract class of Markov transitions given by

$$K_m(x, dy) = \theta(m) \; \delta_x(dy) + (1 - \theta(m)) \; \Phi(m)(dy)$$

where θ denotes a mapping from $\mathcal{P}(E)$ into $[0,1]$ and Φ a mapping from $\mathcal{P}(E)$ into itself. Assume that

$$\sup_{\mathcal{P}(E)} \theta = \|\theta\|_\infty < 1 \quad \text{and} \quad |\theta(m_1) - \theta(m_2)| \leq k(\theta) \; \|m_1 - m_2\|_{tv}$$

for some finite constant $k(\theta) < \infty$ and for any $m_1, m_2 \in \mathcal{P}(E)$. Since we have that

$$K_m(f)(x) - K_m(f)(y) = \theta(m) \; (f(x) - f(y))$$

we readily observe that for any distribution m on E

$$\beta(K_m) = \theta(m) \;\leq\; \|\theta\|_\infty < 1$$

We also notice that $\pi(m) = \Phi(m)$ and

$$\Phi \text{ satisfies } (\mathbf{H}\Phi) \Longrightarrow K_m \text{ satisfies } (9)$$

By induction on the time parameter n it is also easily verified that

$$K_m^n(x, dy) = \theta(m)^n \ \delta_x(dy) + (1 - \theta(m)^n) \ \Phi(m)$$

Furthermore if f_1 and f_2 are chosen such that $\Phi(m_1)(f_1) = 0 = \Phi(m_2)(f_2)$ then we have that

$$\begin{aligned} K_{m_1}^n(f_1) - K_{m_2}^n(f_2) &= \theta(m_1)^n \ f_1 - \theta(m_2)^n \ f_2 \\ &= \theta(m_1)^n \ (f_1 - f_2) + f_2 \ (\theta(m_1)^n - \theta(m_2)^n) \end{aligned}$$

Under our assumptions we also have

$$\begin{aligned} |\theta(m_1)^n - \theta(m_2)^n| &\leq n \ (\theta(m_1) \vee \theta(m_2))^{n-1} \ |\theta(m_1) - \theta(m_2)| \\ &\leq n \ \|\theta\|_\infty^{n-1} |\theta(m_1) - \theta(m_2)| \\ &\leq n \ \|\theta\|_\infty^{n-1} \ k(\theta) \ \|m_1 - m_2\|_{tv} \end{aligned}$$

from which we conclude that

$$\begin{aligned} \|K_{m_1}^n(f_1) - K_{m_2}^n(f_2)\|_\infty &\leq \|\theta\|_\infty^n \ \|f_1 - f_2\|_\infty + n \ \|\theta\|_\infty^{n-1} \ k(\theta) \ \|m_1 - m_2\|_{tv} \\ &\leq n \ \|\theta\|_\infty^{n-1} \ (1 \vee k(\theta)) \ [\|f_1 - f_2\|_\infty + \|m_1 - m_2\|_{tv}] \end{aligned}$$

for any f_1 and f_2 such that $\Phi(m_1)(f_1) = 0 = \Phi(m_2)(f_2)$ and $\|f_1\|_\infty, \|f_2\|_\infty \leq 1$. This implies that condition (10) is satisfied for any class of uniformly bounded functions $\mathcal{F}$ with $\epsilon(n) = n \ \|\theta\|_\infty^{n-1} \ (1 \vee k(\theta))$.
To illustrate this class of models we examine two situations. Consider the Markov transitions

$$K_m(x, dy) = \epsilon \ \delta_x(dy) + (1 - \epsilon) \ mK(dy)$$

where $\epsilon \in [0, 1)$ is a fixed parameter and K a given Markov transition on E. This model clearly belongs to the previous class with $\theta(m) = \epsilon$ and a linear mapping $\Phi(m) = mK$. Now we return to the genetic type SIMC model (2) but we suppose the mutation transition depend on the measure m. More precisely we consider the mapping

$$\Phi'(m)(f) = \frac{m(gK'_m(f))}{m(g)}$$

with

$$K'_m(x, dy) = \theta(m) \ m(dy) + (1 - \theta(m)) \ K(x, dy)$$

where θ denotes a mapping from $\mathcal{P}(E)$ into $[0, 1]$. In this situation we can rewrite $\Phi'(m)$ in the following form

$$\Phi'(m) = mK_m$$

with

$$K_m(x, dy) = \theta(m) \ \delta_x(dy) + (1 - \theta(m)) \ \Phi(m)(dy)$$

where $\Phi(m)$ is the distribution defined in (2).

3.2.3 Genetic self-interactions

In this final section we discuss two different classes of genetic SIMC models associated to a collection of transitions K_m. The first one is the genetic algorithm associated to (7). We suppose that the potential function g is chosen such that $g_{min} > 1$ and K_m is given by

$$K_m(x, dy) = 1/m(g)\ K(x, dy) + (1 - 1/m(g))\ \Phi'(m)(dy) \tag{34}$$

with

$$\Phi'(m)(f) = \frac{m(g'\ K(f))}{m(g')} \quad \text{and} \quad g' = g - 1$$

Our objective is to find conditions on the pair (g, K) underwhich the collection of Markov transitions K_m satisfies $(\mathbf{H}K)$.

The strategy we are going to present is not restricted to this particular genetic model. Next we suppose the transition K_m has the following form

$$K_m(x, dy) = \theta(m)\ K(x, dy) + (1 - \theta(m))\ \Phi'(m)(dy) \tag{35}$$

where $\theta : \mathcal{P}(E) \to (0, 1)$ is a given measurable mapping such that $\|\theta\|_\infty < 1$. This condition ensures that

$$\beta(K_m) = \theta(m)\beta(K) \quad \text{and} \quad \sup_m \beta(K_m) = \|\theta\|_\infty \beta(K) < 1$$

This clearly implies that the mapping π is well defined on $\mathcal{P}(E)$. We also assume that for each $\mu \in \mathcal{P}(E)$ we can find a subset $H^\theta(\mu) \subset \mathrm{Bosc}_1(E)$ such that

$$\left| \frac{1}{1 - \theta(\mu)} - \frac{1}{1 - \theta(\nu)} \right| \leq \sum_{h \in H^\theta(\mu)} b_\mu(h)\ |\mu(h) - \nu(h)| \tag{36}$$

for each $\nu \in \mathcal{P}(E)$ with

$$k(\theta) = \sum_{h \in H^\theta(\mu)} b_\mu(h)\ \in (0, 1)$$

We finally suppose that $\Phi' : \mathcal{P}(E) \to \mathcal{P}(E)$ is a given measurable mapping which satisfies $(\mathbf{H}\Phi)$ for the class of function $\mathcal{F} = \mathrm{Bosc}_1(E)$. By a simple induction argument on the time parameter we prove the following lemma.

Lemma 3.5 *For any $n \geq 1$ and $m \in \mathcal{P}(E)$ we have*

$$K_m^n = \theta(m)^n K^n + (1 - \theta(m)) \sum_{p=0}^{n-1} \theta(m)^p\ \Phi'(m) K^p$$

with the convention $K^0(x, dy) = \delta_x(dy)$ and $\pi(m) = \Phi'(m) R_m$ where R_m is the θ-geometric resolvent Markov transition associated to K and defined by

$$R_m = (1 - \theta(m)) \sum_{n \geq 0} \theta(m)^n\ K^n$$

Then we have the following proposition.

Proposition 3.6 *The collection of transitions K_m defined in (35) satisfies condition* $(\mathbf{H}K)$ *for the class of functions $\mathcal{F} = \mathrm{Bosc}_1(E)$. For any $(f,\mu) \in (\mathcal{F} \times \mathcal{P}(E))$ we can associate a subset $H(f,\mu) \subset \mathcal{F}$ and a collection of numbers $b_{f,\mu}(h) \in [0,1)$, $h \in H(f,\mu)$, such that (9) is met with*

$$\beta_{\mathcal{F}}(\pi) = 2k(\theta) + \beta_{\mathcal{F}}(\Phi')/(1 - \|\theta\|_\infty \beta(K)) \tag{37}$$

In particular the genetic SIMC model (34) satisfies $(\mathbf{H}K)$ *for the same class of functions $\mathcal{F}$ and*

$$\beta_{\mathcal{F}}(\pi) = \frac{2\mathrm{osc}(g)}{(g_{min}-1)^2} + \beta(K)\ \frac{g_{min}\ (2\mathrm{osc}(g) + (g_{min}-1))}{(g_{min} - \beta(K))\ (g_{min}-1)}$$

and $\beta_{\mathcal{F}}(\pi) < 1$ as soon as the pair (g,K) satisfies

$$\beta(K) < \frac{1}{g_{min}}\ \frac{(g_{min}-1)^2 - 2\mathrm{osc}(g)}{(g_{min}-1) + 2\mathrm{osc}(g)} \tag{38}$$

Proof: To prove that π has the desired regularity properties we use the decomposition

$$\pi(m_1) - \pi(m_2) = \Phi'(m_1)(R_{m_1} - R_{m_2}) + (\Phi'(m_1) - \Phi'(m_2))R_{m_2} \tag{39}$$

For each $f \in \mathcal{F}$ we have

$$\begin{aligned}
&(\Phi'(m_1) - \Phi'(m_2))R_{m_2}(f) \\
&= (1-\theta(m_2)) \sum_{n\geq 0} \theta(m_2)^n\ (\Phi'(m_1) - \Phi'(m_2))K^n(f) \\
&= (1-\theta(m_2)) \sum_{n\geq 0} \theta(m_2)^n\ \beta(K)^n\ (\Phi'(m_1)(f^{(n)}_{m_2}) - \Phi'(m_2)(f^{(n)}_{m_2}))
\end{aligned}$$

with

$$f^{(n)}_{m_2} = [K^n(f) - \Phi'(m_2)K^n(f)]/\beta(K)^n \quad (\in \mathcal{F})$$

Under our assumptions for each $n \geq 0$, $f \in \mathcal{F}$ and $m_2 \in \mathcal{P}(E)$ we can find a subset $H^{(n)}(m_2,f) \subset \mathcal{F}$ and a collection of numbers $b^{(n)}_{m_2,f}(h)$, $h \in H^{(n)}(m_2,f)$ such that

$$|\Phi'(m_1)(f^{(n)}_{m_2}) - \Phi'(m_2)(f^{(n)}_{m_2})| \leq \sum_{h\in H^{(n)}(m_2,f)} b^{(n)}_{m_2,f}(h)\ |m_1(h) - m_2(h)|$$

It follows that

$$\begin{aligned}
&|\Phi'(m_1) - \Phi'(m_2))R_{m_2}(f)| \\
&\leq \textstyle\sum_{n\geq 0} \sum_{h\in H^{(n)}(m_2,f)} (\|\theta\|_\infty \beta(K))^n\ b^{(n)}_{m_2,f}(h)\ |m_1(h) - m_2(h)|
\end{aligned} \tag{40}$$

Before we continue we notice that

$$\sup_{f\in\mathcal{F}, m_2\in\mathcal{P}(E)} \sum_{n\geq 0} \sum_{h\in H(m_2,f)} (\|\theta\|_\infty \beta(K))^n\ b_{m_2,f}(h)\ = \frac{\beta_{\mathcal{F}}(\Phi')}{1 - \|\theta\|_\infty \beta(K)}$$

To treat the first term in (39) we use the formula

$$\begin{aligned}
&\Phi'(m_1)(R_{m_1} - R_{m_2})(f) \\
&= \sum_{n\geq 0}[(1-\theta(m_1))\theta(m_1)^n - (1-\theta(m_2))\theta(m_2)^n]\ \Phi'(m_1)K^n(f) \\
&= \sum_{n\geq 0}[(\theta(m_1)^n - \theta(m_2)^n) + (\theta(m_2)^{n+1} - \theta(m_1)^{n+1})]\ \Phi'(m_1)K^n(f)
\end{aligned}$$

According to this representation we find that

$$|\Phi'(m_1)(R_{m_1} - R_{m_2})(f)| \leq \sum_{n\geq 0} |\theta(m_1)^n - \theta(m_2)^n| + \sum_{n\geq 0} |\theta(m_2)^{n+1} - \theta(m_1)^{n+1}|$$

If we define

$$\theta_1 = \theta(m_1) \vee \theta(m_2) \quad \text{and} \quad \theta_2 = \theta(m_1) \wedge \theta(m_2)$$

we check that

$$\begin{aligned}
|\Phi'(m_1)(R_{m_1} - R_{m_2})(f)| &\leq \frac{1}{1-\theta_1} - \frac{1}{1-\theta_2} + \frac{\theta_1}{1-\theta_1} - \frac{\theta_2}{1-\theta_2} \\
&= 2\left(\frac{1}{1-\theta_1} - \frac{1}{1-\theta_2}\right)
\end{aligned}$$

Under our assumptions this implies that

$$|\Phi'(m_1)(R_{m_1} - R_{m_2})(f)| \leq 2 \sum_{h\in H^\theta(m_2)} b_{m_2}(h)\ |m_1(h) - m_2(h)| \tag{41}$$

If we combine (40) and (41) we find that the mapping π has the desired regularity. Ne now come to the proof of (10). We start by noting that

$$\pi(m)(f) = 0 \Rightarrow \sum_{p=0}^{n-1} \theta(m)^p\ \Phi'(m)K^p(f) = -\sum_{p\geq n} \theta(m)^p\ \Phi'(m)K^p(f)$$

Let $(m_1, m_2) \in (\mathcal{P}(E) \times \mathcal{P}(E))$ and let $(f_1, f_2) \in \mathcal{F}$ be a pair of functions such that $\pi(m_1)(f_1) = 0 = \pi(m_2)(f_2)$. From the above observation we have

$$\begin{aligned}
&K^n_{m_1}(f_1) - K^n_{m_2}(f_2) \\
&= \theta(m_1)^n K^n(f_1) - \theta(m_2)^n K^n(f_2) \\
&+ (1-\theta(m_2))\sum_{p\geq n} \theta(m_2)^p\ \Phi'(m_2)K^p(f_2) - (1-\theta(m_1))\sum_{p\geq n}\theta(m_1)^p\ \Phi'(m_1)K^p(f_1)
\end{aligned}$$

To estimate this difference we introduce the decomposition

$$K^n_{m_1}(f_1) - K^n_{m_2}(f_2) = I_1 + I_2 + I_3 + I_4 + I_5$$

with

$$\begin{aligned}
I_1 &= [\theta(m_1)^n - \theta(m_2)^n]\ K^n(f_1) \\
I_2 &= \theta(m_2)^n\ [K^n(f_1) - K^n(f_2)]
\end{aligned}$$

$$
\begin{aligned}
I_3 &= (1-\theta(m_2)) \sum_{p\geq n} \theta(m_2)^p \; [\Phi'(m_2) - \Phi'(m_1)] K^p(f_2) \\
I_4 &= (1-\theta(m_2)) \sum_{p\geq n} \theta(m_2)^p \; \Phi'(m_1) [K^p(f_2) - K^p(f_1)] \\
I_5 &= \sum_{p\geq n} [(1-\theta(m_2))\theta(m_2)^p - (1-\theta(m_1))\theta(m_1)^p] \; \Phi'(m_1) K^p(f_1)
\end{aligned}
$$

Under our assumptions we have

$$
\begin{aligned}
|I_1| &\leq |\theta(m_1)^n - \theta(m_2)^n| \leq n \; \|\theta\|_\infty^{n-1} \sum_{h\in H^\theta(m_2)} b_{m_2}(h) \; |m_1(h) - m_2(h)| \\
&\leq n \; \|\theta\|_\infty^{n-1} \; k(\theta) \; \|m_1 - m_2\|_{tv}
\end{aligned}
$$

and $|I_2| \leq \|\theta\|_\infty^n \|f_1 - f_2\|_\infty$. Since the mapping Φ' satisfies **(HΦ)** for the class of functions f such that $\mathrm{osc}(f) \vee \|f\|_\infty \leq 1$ we have that

$$
\|\Phi'(m_1) - \Phi'(m_2)\|_{tv} \leq c \; \|m_1 - m_2\|_{tv}
$$

for some finite constant $c < \infty$. This yields that

$$
|I_3| \leq c(1-\theta(m_2)) \sum_{p\geq n} \theta(m_2)^p \; \|m_1 - m_2\|_{tv} \leq c \; \|\theta\|_\infty^n \; \|m_1 - m_2\|_{tv}
$$

We also notice that

$$
|I_4| \leq \frac{\|\theta\|_\infty^n}{1 - \|\theta\|_\infty} \; \|f_1 - f_2\|_\infty
$$

We now evaluate the term I_5. For each pair of numbers $a_1, a_2 \in (0,1)$, $a_1 < a_2$, we have

$$
U_n(a) = |\sum_{p\geq n} (1-a_1) a_1^p - \sum_{p\geq n} (1-a_2) a_2^p| = |a_1^n - a_2^n| \leq n a_2^{n-1} (a_2 - a_1) \tag{42}
$$

Then we note that

$$
a_2 - a_1 = (1-a_1)(1-a_2) \; [\frac{1}{1-a_2} - \frac{1}{1-a_1}]
$$

This yields

$$
U_n(a) \leq n a_2^{n-1} [\frac{1}{1-a_2} - \frac{1}{1-a_1}]
$$

and the end of the proof of (42) is clear. Using this estimate and under our assumptions on the mapping θ we find that

$$
|I_5| \leq 2 \; (n+1) \; k(\theta) \; \|\theta\|_\infty^{n-1} \; \|m_1 - m_2\|_{tv}
$$

The end of the proof of (10) is now a simple combination of the above estimates. Now we examine the genetic type Markov transition (34). If we take $\theta(m) = 1/m(g)$ we clearly have $\|\theta\|_\infty = 1/g_{min}$. Furthermore we notice that

$$
\begin{aligned}
\left| \frac{1}{1-\theta(\mu)} - \frac{1}{1-\theta(\nu)} \right| &= \left| \frac{\mu(g)}{\mu(g-1)} - \frac{\nu(g)}{\nu(g-1)} \right| \\
&= \frac{1}{\mu(g-1)\nu(g-1)} \; |\mu(g) - \nu(g)| \\
&\leq \frac{\mathrm{osc}(g)}{(g_{min}-1)^2} \; \left| \nu\left(\frac{1}{\mathrm{osc}(g)} (g - \mu(g)) \right) \right|
\end{aligned}
$$

It follows that (36) is met with

$$H^\theta(\mu) = \left\{ \frac{1}{\text{osc}(g)}(g - \mu(g)) \right\} \quad \text{and} \quad k(\theta) = \frac{\text{osc}(g)}{(g_{min} - 1)^2}$$

Note that $2k(\theta) < 1$ as soon as $\|g\|_\infty \leq (1 + g_{min}^2)/2$. From (29) we also notice that Φ' satisfies **(HΦ)** for the class of functions $\mathcal{F} = \text{Bosc}_1(E)$

$$\beta_{\mathcal{F}}(\Phi') = \frac{\beta(K)}{g_{min} - 1} \left(\text{osc}(g) + \|g - 1\|_\infty\right)$$

We also note that $\beta_{\mathcal{F}}(\Phi') < 1$ as soon as

$$\sup_{x,y} \frac{g(x) - 1}{g(y) - 1} < \frac{1}{2} \left(\frac{1}{\beta(K)} + 1 \right)$$

If follows from (37) that

$$\beta_{\mathcal{F}}(\pi) = \frac{2\text{osc}(g)}{(g_{min} - 1)^2} + \beta(K) \; \frac{g_{min}(\text{osc}(g) + \|g - 1\|_\infty)}{(g_{min} - \beta(K))(g_{min} - 1)}$$

Since $\|g - 1\|_\infty = \text{osc}(g) + (g_{min} - 1)$ and $(g_{min} - \beta(K)) \geq (g_{min} - 1)$ we obtain

$$\beta_{\mathcal{F}}(\pi) \leq \frac{1}{(g_{min} - 1)^2}[2\text{osc}(g) + \beta(K) \; g_{min}(2\text{osc}(g) + (g_{min} - 1))]$$

The right hand side is strictly less than one as soon as

$$\beta(K) < \frac{1}{g_{min}} \left[\frac{(g_{min} - 1)^2 - 2\text{osc}(g)}{(g_{min} - 1) + 2\text{osc}(g)} \right]$$

This completes the proof of the proposition. ■

We end this section with the genetic type SIMC model (8). We recall that the Markov transitions are defined by

$$K_m(x, dy) \;=\; a\; K(x, dy) \;+\; b\; \Psi(m)(dy) \quad \text{with} \quad \Psi(m)(f) = m(fg)/m(g)$$

for some $a, b \in [0, 1]$, $a + b = 1$, a positive fitness function g and a Markov transition K.

Proposition 3.7 *Suppose the triplet (a, g, K) satisfies the following condition*

$$a\; \beta(K) < 1\;, \quad g_{min} = 1 \quad \text{and} \quad \text{osc}(g) < \frac{a}{2b}\; (1 - \beta(K))$$

For any $m \in \mathcal{P}(E)$ and $n \geq 1$ we have

$$K_m^n \;=\; a^n\; K^n + b \sum_{p=0}^{n-1} a^p\; \Psi(m)K^p \tag{43}$$

$$\pi(m) \;=\; \Psi(m)K_a \quad \text{with} \quad K_a = b\sum_{n\geq 0} a^n K^n \tag{44}$$

Furthermore the Markov transitions K_m satisfy **(HK)** *with*

$$\mathcal{F} = \text{Bosc}_1(E) \quad \text{and} \quad \beta_{\mathcal{F}}(\pi) = (1 + 2\text{osc}(g))\; \frac{b}{1 - \beta(K)a}$$

Proof: Proceeding inductively on the time parameter we prove (43). Since

$$K_m(f)(x) - K_m(f)(y) \;=\; a\ (K(f)(x) - K(f)(y))$$

we clearly have that

$$\beta(K_m) = a\ \beta(K)\ (< 1)$$

Thus, the mapping π is well defined and it is easily checked that it is given by (44). Now we observe that for each $f \in \mathrm{Bosc}_1(E)$ and any pair of distributions $(m_1, m_2) \in \mathcal{P}(E) \times \mathcal{P}(E)$

$$\begin{aligned}\pi(m_1)(f) - \pi(m_2)(f) &= b \sum_{n\geq 0} a^n\ [\Psi(m_1)K^n(f) - \Psi(m_2)K^n(f)] \\ &= b \sum_{n\geq 0} a^n\ \frac{1}{m_1(g)}\ [m_1(f^{(n)}_{m_2}) - m_2(f^{(n)}_{m_2})]\end{aligned}$$

with

$$f^{(n)}_{m_2} \;=\; g\ K^n(f - \Psi(m_2)K^n(f))$$

Arguing as in section 3.1 we can check that

$$|f^{(n)}_{m_2}(x) - f^{(n)}_{m_2}(y)| \;\leq\; \beta(K)^n\ (1 + 2\mathrm{osc}(g))\ \mathrm{osc}(f)$$

and

$$\|f^{(n)}_{m_2}\|_\infty \;\leq\; \beta(K)^n\ (1 + \mathrm{osc}(g))\ \mathrm{osc}(f)$$

If we write

$$\overline{f}^{(n)}_\mu(x) = \frac{1}{\beta(K)^n\ (1 + 2\mathrm{osc}(g))}\ f^{(n)}_\mu(x)$$

then we conclude that (9) is satisfied with

$$\begin{aligned}\mathcal{F} &= \{f \in \mathrm{Osc}_1(E)\ ;\ \|f\|_\infty \leq 1\} \\ H(f,\mu) &= \{\ \overline{f}^{(n)}_\mu\ , n \geq 0\ \ \} \\ b_{f,\mu}(h^{(n)}_\mu) &= b\ (1 + 2\mathrm{osc}(g))\ (a\ \beta(K))^n\end{aligned}$$

Summing $b_{f,\mu}(h^{(n)}_\mu)$ from $n = 0$ to ∞ gives

$$\beta_{\mathcal{F}}(\pi) = (1 + 2\mathrm{osc}(g))\ \frac{b}{1 - \beta(K)a}\ (< 1)$$

Let us check (10). First we note that if $\pi(m)(f) = 0$ then

$$b \sum_{p=0}^{n-1} a^p\ \Psi(m)K^p(f) = -b \sum_{p\geq n} a^p\ \Psi(m)K^p(f)$$

From this observation and for any pair of functions f_1, f_2 such that

$$\pi(m_1)(f_1) = 0 = \pi(m_2)(f_2)$$

using classical arguments we have that there exists a constant $c \geq 0$ such that

$$\begin{aligned}\|K^n_{m_1}(f_1) - K^n_{m_2}(f_2)\|_\infty &\leq a^n\ \|K^n(f_1) - K^n(f_2)\|_\infty \\ &\quad + \sum_{p\geq n} ba^p\ |\Psi(m_2)K^p(f_2) - \Psi(m_1)K^p(f_2)|\end{aligned}$$

$$
\begin{aligned}
&\quad + \sum_{p \geq n} ba^p \; |\Psi(m_1) K^p (f_2 - f_1)| \\
\leq \quad & a^n \; \|f_1 - f_2\|_\infty + c \sum_{p \geq n} ba^p \; \|m_1 - m_2\|_{tv} \\
&\quad + \sum_{p \geq n} ba^p \; \|f_2 - f_1\|_\infty
\end{aligned}
$$

The above estimate implies that

$$\|K^n_{m_1}(f_1) - K^n_{m_2}(f_2)\|_\infty \leq a^n (2 \vee c) \; [\|f_2 - f_1\|_\infty + \|m_1 - m_2\|_{tv}]$$

and (10) clearly holds with $\epsilon(n) = (2 \vee c) \; a^n$. ■

Remark 3.8: Let $h : E \to (0, \infty)$ be a potential function such that $h_{min} > 0$ and let $K^{(h)}_m$ be the Markov transition defined as above by replacing g by h. More precisely $K^{(h)}_m$ is given by

$$K^{(h)}_m(x, dy) \; = \; a \, K(x, dy) + b \, \Psi^{(h)}(m)(dy) \quad \text{with} \quad \Psi^{(h)}(m)(f) = m(fh)/m(h)$$

Also denote by $\pi^{(h)}(m)$ the mapping defined as π by replacing in (44) Ψ by $\Psi^{(h)}$. Arguing as above we find that $K^{(h)}_m$ satisfies $(\mathbf{H}K)$ with the same class of function $\mathcal{F}$ and

$$\beta_{\mathcal{F}}(\pi^{(h)}) = \frac{b}{h_{min}} \; \frac{(\text{osc}(h) + \|h\|_\infty)}{1 - \beta(K) a}$$

Furthermore we have $\beta_{\mathcal{F}}(\pi^{(h)}) < 1$ as soon as

$$\sup_{x,y} \frac{h(x)}{h(y)} < 1 + \frac{a}{2b} \; (1 - \beta(K))$$

We now consider the SIMC model (1) associated to the mapping

$$\Phi(m) = \pi(m) = \Psi(m) K_a$$

Observe that the resulting genetic SIMC model coincide with the genetic model (2) with mutation transition K_a. We also notice that K_a is the a-geometric resolvent kernel of the Markov transition K and

$$\beta(K_a) \leq b/(1 - a\beta(K))$$

Probabilistically K_a is the Markov kernel of a chain “sampled” at geometric time points. More precisely the mutation is decomposed in two step. First we select a time parameter $p \in \mathbb{N}$ with probability $b \; a^p$. Then the particle evolves according to the transition K^p. Roughly speaking for large values of the parameter a this mutation mechanism is more mixing than K. More interestingly we can initiate a comparison of the decays to equilibrium of the two genetic SIMC models (2) and (8). We start by noting that the corresponding coefficients $\beta_{\mathcal{F}}(\Phi)$ and $\beta_{\mathcal{F}}(\pi)$ given in (4) and (11) are related by the formula

$$\frac{\beta_{\mathcal{F}}(\Phi)}{\beta_{\mathcal{F}}(\pi)} = \frac{1}{b} \; \beta(K) \; (1 - a\beta(K))$$

When $2a \leq 1$ the mapping $\beta \in [0, 1] \to u(\beta) = \beta \; (1 - a\beta)/b$ is increasing from $u(0) = 0$ to $u(1) = 1$ and we always have that $\beta_{\mathcal{F}}(\Phi) \leq \beta_{\mathcal{F}}(\pi)$.

When $2a > 1$ the mapping u increases on $[0, 1/(2a)]$ from $u(0) = 0$ to $u(1/2a) = 1/(4ab)$ and it decreases on $[1/(2a), 1]$ from $u(1/2a) = 1/(4ab)$ to $u(1) = 1$. Noting that $u(b/a) = 1$ we also have that

$$\beta_{\mathcal{F}}(\pi) \leq \beta_{\mathcal{F}}(\Phi) \iff \beta(K) \geq b/a$$

It is also a simple exercise to check for instance that if $ab \leq 1/7$ then

$$u(\beta) = 7/4 \iff \beta = \frac{7}{12}\,(1 - 1/\sqrt{7}) \quad \text{or} \quad \beta = \frac{7}{12}\,(1 + 1/\sqrt{7})$$

For instance if we take $a = 6/7$ and $\beta(K) = 7/12\ (1 + 1/\sqrt{7})$ then we have that

$$\beta_{\mathcal{F}}(\Phi) = \frac{7}{4}\,\beta_{\mathcal{F}}(\pi)$$

When the fitness function is chosen accordingly we may have that

$$\beta_{\mathcal{F}}(\pi) = 1/2 < \beta_{\mathcal{F}}(\Phi) = 7/8$$

In this situation the decays to equilibrium for the SIMC model (8) and (2) are respectively $n^{1/2}$ and $n^{1/4}$. In other words if we use more mutations in the SIMC model (8) theorem 1.2 leads to faster convergence.

References

[1] M. Benaim. Vertex-reinforced random walks and a conjecture of Pemantle. Ann. Probab. vol. 25, n. 1, pp. 361–392, 1997.

[2] M. Benaim, M. Ledoux, O. Raimond. Self-interacting diffusions Probab Theory Relat Fields 122, 1, 1-41, (2002)

[3] D. Coppersmith, P. Diaconis. Random walks with reinforcement, Technical report, Standford University 1987.

[4] P. Del Moral. *Measure valued processes and interacting particle systems. Application to non-linear filtering problems*, Ann. Applied Probab. vol. 8, n. 2, pp. 438–495, 1998.

[5] P. Del Moral and A. Guionnet. *On the stability of interacting processes with applications to filtering and genetic algorithms.* Ann. Inst. H. Poincaré, vol. 37, n. 2, pp. 155-194, 2001.

[6] P. Del Moral and M.A. Kouritzin and L. Miclo. *On a class of discrete generation interacting particle systems*, Electronic Journal of Probability, vol. 6, n. 16, pp. 1–26, 2001.

[7] P. Del Moral and L. Miclo, *Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae with Applications to Non-Linear Filtering*, Séminaire de Probabilités XXXIV, Springer-Verlag, Berlin, J. Azéma, M. Emery, M. Ledoux and M. Yor editors, Lecture Notes in Mathematics 1729, pp. 1–145, 2000.

[8] P. Del Moral and L. Miclo. *Particle approximations of Lyapunov exponents connected to Feynman-Kac-Scrhödinger semigroups.* Publications du Laboratoire de Statistique et Probabilités no 01-08, 2001. Available at http://www.lsp.ups-tlse.fr/Delmoral/preprints.html.

[9] P. Del Moral and L. Miclo. *On the Convergence and the Applications of the Generalized Simulated Annealing.* SIAM Journal on Control and Optimization, vol. 37, n. 4, pp. 1222-1250, 1999.

[10] B. Davis. *Reinforced random walk.* Probab. Theory and Related Fields, vol. 84, n. 2, pp. 203-229, 1990.

[11] R.L. Dobrushin. Central limit theorem for nonstationnary Markov chains, I,II, Theory of Probability and its applications, vol.1, n. 1 and 4, 1956, pp. 66-80 and pp. 330-385.

[12] P. Donnelly. *Partition structures, Polya urns, Ewens sampling formula and the ages of alleles.* Theoret. Population Biol. vol. 30, n. 2, pp. 271–288, 1986.

[13] P.J. Flory. *Principles of polymer chemistry*, Cornell University Press, Ithaca, New York, 1962.

[14] F.M. Hoppe. *The sampling theory of neutral alleles and an urn model in population genetics.* J. Math. Biol., vol. 25, n. 2, pp. 123–159, 1987.

[15] R. Pemantle. *Phase transition in reinforced random walk and RDWRE on trees*, Ann. Probab. vol. 16, n. 3, pp.1229-1241, 1988.

[16] R. Pemantle. *Vertex-reinforced random walk.* Probab. Theory Related Fields, vol. 92, n. 1, pp. 117–136, 1992.

[17] R. Pemantle, S Volkov. *Vertex-reinforced random walk on $\mathbb{Z}$ has finite range*, Ann. Probab. vol. 27, n. 3, pp. 1368–1388, 1999.

[18] E. Renshaw. *Computer simulation of Stitka spruce: spatial branching models for canopy growth and root structure.* IMA J. Math. Appl. Med. Biol., vol. 2, pp. 183–200, 1985.

[19] T. Sellke. *Reinforced random walk on the d-dimensional integer lattice.* Technical report Purdue University 1994.

[20] A.N. Shiryaev. Probability, *Graduate texts in Mathematics* 95, Second edition, Springer 1996.

[21] G. Trieb. *A Polya urn model and the coalescent*, J. Applied Probab. vol. 29, n. 1, pp. 1–10, 1992.

[22] B. Tóth. *limit theorems for weakly reinforced random walks on $\mathbb{Z}$*, Studia Sci. Math. Hungar. vol. 33, n. 1-3, pp. 321-337, 1997.