Extremality of invariant measures and ergodicity of stochastic systems

2002 J. Phys. A: Math. Gen. 35 229

(http://iopscience.iop.org/0305-4470/35/2/304)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 229–238 PII: S0305-4470(02)23450-2

Extremality of invariant measures and ergodicity of stochastic systems

Jinwen Chen

Department of Mathematical Sciences, Tsinghua University, Beijing 100084, People's Republic of China

E-mail: jchen@math.tsinghua.edu.cn

Received 27 March 2001, in final form 13 September 2001
Published 4 January 2002
Online at stacks.iop.org/JPhysA/35/229

Abstract
In this paper we first present some conditions on the extremality of invariant measures of a Markov process. These conditions involve average convergence and invariant functions of the process. We then combine these with coupling and duality to study some aspects of convergence to invariant measures of Markov processes. In particular, we suggest a new way of studying complete convergence. Applications to interacting particle systems are given.

PACS numbers: 02.50.Ga, 02.50.Ey, 05.10.Gg, 05.50.+q

1. Introduction

Investigating invariant measures is one of the most important methods for studying a stochastically interacting system which evolves in time. Usually, in physics or other fields, these measures are used to describe the equilibrium or stable states of such a system. Thus better characterizations of these invariant measures are more helpful, and even crucial in many cases, in understanding the asymptotic behaviour, such as ergodicity, metastability and fluctuation, etc, of the system. Typically, this can be reduced to characterizing the extremal invariant measures. For example, for many important models, such as attractive spin flip systems, the ergodicity is equivalent to the uniqueness of such extremal invariant measures, see [9]. The motivation for considering the problems in this paper comes from the study of interacting particle systems (IPS), where one of the most important problems is to determine the set of all the invariant measures of a system under consideration, and then to describe the domain of attraction for each of these invariant measures, i.e. starting from various initial configurations to find out which given equilibrium state the system will approach as time goes to infinity. In this paper we first try to describe the set of the invariant measures by characterizing those extremal invariant measures from various aspects. Then we combine these results with some powerful techniques, such as coupling and duality, etc, to study other problems related to ergodicity, including convergence to invariant measures and complete convergence (which specifies the limiting law for each initial configuration), etc, the former is related to some interesting open problems in IPS (see corollary 4.2 and [9]). We first work with abstract Markov processes to obtain some general results. Then we apply these results to IPS.

There are different notions of ergodicity, all involving the properties of the invariant probability measures (stationary distributions) and certain convergence of the corresponding Markov processes. We will specify the most commonly considered one in IPS and some related modifications; see the definition below. Let E be a Polish space, $\Omega = D([0, \infty), E)$ be the space of functions from $[0, \infty)$ to E which are right continuous with left limits, equipped with the Skorohod topology. $\{P_x, x \in E\}$ is assumed to be a Markov family of probability measures on Ω, with P_x being weakly continuous in x. $\{S(t), t \geqslant 0\}$ denotes the corresponding Markov semigroup. Denote by $m_1(E)$ the set of probability measures on E. For $\mu \in m_1(E)$ and $t \geqslant 0$, $\mu S(t) \in m_1(E)$ is defined by

$$\int f \, \mathrm{d}\mu S(t) = \int S(t) f \, \mathrm{d}\mu,$$

which is the law of the process at time t, starting with initial law μ. μ is said to be invariant (or stationary) for $\{S(t), t \geqslant 0\}$ (or for the Markov process) if $\mu S(t) = \mu \, \forall t \geqslant 0$. Denote by $m_i(E)$ the set of all such invariant μ.

Definition.

(1) A Markov process on Ω with semigroup $\{S(t), t \geqslant 0\}$ is said to be ergodic, if $m_i(E) = \{\nu\}$ is a singleton and for each $\mu \in m_1(E)$, $\lim_{t\to\infty} \mu S(t) = \nu$ weakly. The ergodicity is said to be a strong one if the weak convergence is replaced by the convergence in the τ-topology, i.e. for every measurable subset A of E,

$$\lim_{t\to\infty} \mu S(t)(A) = \nu(A).$$

The ergodicity is said to be uniform if the convergence holds in the total variation norm $\|\cdot\|$. Here we recall that the total variation $\|\mu\|$ of a signed measure μ is equivalently defined by

$$\|\mu\| = \sup\left\{ \int f \, \mathrm{d}\mu \colon f \text{ bounded measurable with } \|f\|_\infty \equiv \sup|f| \leqslant 1 \right\}.$$

(2) If we replace $S(t)$ by $T(t) = \frac{1}{t}\int_0^t S(u)\,\mathrm{d}u$, then the corresponding ergodicity will be said to be average.

Thus to understand the ergodic behaviour, the first key step is to find out if $m_i(E)$ is a singleton. It is known that this can be reduced to study if there is a unique extremal stationary distribution (see, e.g., [9]). Therefore it is interesting to describe the set of all extremal measures in $m_i(E)$, denoted by $m_i^e(E)$, as clearly as possible. In this paper we will present in section 2 some equivalent conditions for extremality of $\mu \in m_i(E)$ (theorem 2.1), seen from various points of view. As an example of applications of this result, in section 4 we show extremality of a class of invariant measures for certain spin flip systems. These equivalent conditions concern average convergence and invariant functions of the semigroup. These notions are closely related to coupling and ergodicity. Consequently we then discuss ergodicity in terms of average convergence and extremal invariant measures. In particular, our results imply that strong average convergence is equivalent to the uniform one (theorem 2.2).

In section 3 we apply our results concerning extremality to study some aspects of convergence to invariant measures, especially the so-called complete convergence, of Markov processes. Combining these with a duality argument, in section 4 we discuss some potential applications of our results to interacting particle systems. We show that for a large class of typical systems, the complete convergence consideration can be reduced to studying some fundamental problems for certain continuous time Markov chains—a type of simpler and well studied Markov process. We would like to point out that studying complete convergence is usually hard work. For complete convergence of IPS, most of the techniques used in previous works heavily rely on the specific construction of the individual models under consideration (see, e.g., [9]). Our approach may be regarded as an attempt at exploring new and effective, and universal in some sense, ways for studying this problem. In particular, as an example we show that every linear voter or anti-voter model is completely convergent (see example 1 in section 4). We hope that this approach can be further developed into a more practical and effective technique. Our duality arguments also enable us to obtain a result on convergence to invariant measures of these systems, which partially answered some important open problem proposed in [9] for IPS.

2. Extremality of the invariant measure

Recall the definitions of $\mathbf{S} = \{S(t),\ t \geqslant 0\}$, $\{T(t),\ t \geqslant 0\}$, $m_1(E), m_i(E)$, and $m_i^e(E)$. For each $\nu \in E_i(E)$, $\mathbf{S}$ can be uniquely extended to $L^2(\nu)$, denoted by $\mathbf{S}^\nu = \{S^\nu(t),\ t \geqslant 0\}$. A function $f \in L^2(\nu)$ is called $\mathbf{S}^\nu$-*invariant*, if $\mathbf{S}^\nu(t)f = f$, $\forall t \geqslant 0$ ν-a.s. We denote $T^\nu(t)f = \frac{1}{t}\int_0^t S^\nu(u)f\,\mathrm{d}u$. $m_{i,\nu}(E)$ is the set of measures in $m_i(E)$ that are absolutely continuous w.r.t. ν. Then we have the following theorem:

Theorem 1. *If $m_i(E) \neq \emptyset$, then $m_i^e(E) \neq \emptyset$, and for $\nu \in m_i(E)$ the following assertions are equivalent:*

(1) $\nu \in m_i^e(E)$;
(2) $\forall f \in L^2(\nu)$, $\lim_{t\to\infty} T^\nu(t)f = \nu(f)$ in $L^2(\nu)$;
(3) For $f \in L^2(\nu)$, if for some sequence $t_n \to +\infty$, $T^\nu(t_n)f = f$ ν-a.s., then $f = const$ ν-a.s.;
(4) Every $\mathbf{S}^\nu$-invariant function in $L^2(\nu)$ is ν-a.s. constant;
(5) Every bounded $\mathbf{S}^\nu$-invariant function is ν-a.s. constant;
(6) $m_{i,\nu} = \{\nu\}$.

Furthermore, different measures in $m_i^e(E)$ are singular w.r.t. each other.

Proof. For $\nu \in m_1(E)$, let $P_\nu = \int P_x\nu(\mathrm{d}x)$. From [4], theorem 1, P_ν can be uniquely represented as a generalized convex combination of stationary ergodic Markov measures on Ω, and a stationary and ergodic measure Q on Ω is Markov associated with S iff its single time marginal μ_Q is in $m_i^e(E)$. Therefore, $m_i(E) \neq \emptyset$ implies $m_i^e(E) \neq \emptyset$.

(1) $\Rightarrow$ (2). Let $\nu \in m_i^e(E)$, then P_ν is ergodic as just stated. Thus by the well-known ergodic theorem, if f is bounded measurable on E, then

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(x_s)\,\mathrm{d}s = \nu(f) = \left(\int f\,\mathrm{d}\nu\right) \qquad P_\nu\text{-a.s. and in } L^1(P_\nu).$$

Hence

$$\int |T^\nu(t)f - \nu(f)|^2\,\mathrm{d}\nu = \int \left|E^{P_x}\frac{1}{t}\int_0^t f(x_u)\,\mathrm{d}u - \nu(f)\right|^2 \mathrm{d}\nu$$
$$\leqslant \int \left|\frac{1}{t}\int_0^t f(x_u)\,\mathrm{d}u - \nu(f)\right|^2 \mathrm{d}P_\nu \to 0 \qquad (t\to\infty)$$

The extension to $f \in L^2(\nu)$ is standard if one uses the fact that $C_b(E)$ is dense in $L^2(\nu)$.

Obviously (2) $\Rightarrow$ (3), (3) $\Rightarrow$ (4) and (4) $\Rightarrow$ (5). The proof of (2) $\Rightarrow$ (1) is standard. We now prove (5) $\Rightarrow$ (2). Let f be bounded measurable. Since P_ν is stationary on Ω, by the ergodic theorem, the limits

$$g(x.) = \lim_{t\to\infty} \frac{1}{t}\int_0^t f(x_u)\,\mathrm{d}u$$

exist P_ν-a.s. and in $L^1(P_\nu)$. Then g is P_ν-a.s. shift invariant, i.e. $g(x_{t+\cdot}) = g(x.)$ P_ν-a.s., and $\int g\,\mathrm{d}P_\nu = \nu(f)$. Define $g_0(x) = E^{P_x} g$, then

$$\int |S^\nu(t)g_0 - g_0|\,\mathrm{d}\nu = \int \left|E^{P_x}E^{P_{x_t}}g - E^{P_x}g\right|\mathrm{d}\nu.$$

From the Markov property, the right-hand side is equal to

$$\int \left|E^{P_x}[g(x_{t+\cdot}) - g]\right|\mathrm{d}\nu \leqslant \int |g(x_{t+\cdot}) - g(x.)|\,\mathrm{d}P_\nu = 0.$$

This means that g_0 is $\mathbf{S}^\nu$-invariant. By assumption (4), $g_0 =$ constant ν-a.s. Hence

$$\int |T^\nu(t)f - \nu(f)|^2\,\mathrm{d}\nu = \int \left|E^{P_x}\frac{1}{t}\int_0^t f(x_u)\,\mathrm{d}u - E^{P_x}g\right|^2 \mathrm{d}\nu$$

$$\leqslant 2\|f\|_\infty \int \left|\frac{1}{t}\int_0^t f(x_u)\,\mathrm{d}u - g\right|\mathrm{d}P_\nu \to 0 \qquad (t\to\infty)$$

where $\|f\|_\infty$ is the uniform norm of f. Now it remains to prove (1) $\Rightarrow$ (6), since (6) $\Rightarrow$ (1) is trivial. Let $\nu \in m_i^e(E)$ and $\mu \in m_{i,\nu}$. Denote $p = \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$. Then clearly $1 \geqslant \alpha = \int p \wedge 1\,\mathrm{d}\nu > 0$. We now show that indeed $\alpha = 1$. If $\alpha > 1$, define

$$\mathrm{d}\mu_1 = \alpha^{-1}(p\wedge 1)\,\mathrm{d}\nu \qquad \text{and} \qquad \mathrm{d}\mu_2 = (1-\alpha)^{-1}(1 - p\wedge 1)\,\mathrm{d}\nu$$

then from proposition 2.2 in [1] we know that $\mu_1, \mu_2 \in m_i(E)$ and

$$\nu = \alpha\mu_1 + (1-\alpha)\mu_2$$

By the extremality of ν, $\mu_1 = \mu_2$. This implies $p \wedge 1 = \alpha < 1$, ν-a.s. That is, $p < 1$ ν-a.s., contradicting the definition of p.

Thus we obtain that $p = 1$ ν-a.s., i.e. $\mu = \nu$. Hence we have $m_{i,\nu} = \{\nu\}$. The last conclusion of the theorem is a simple consequence of (2). □

Remark 2.1.

(1) We may compare the above theorem with the results in [9], chapter IV, where it was proved that for a stochastical Ising model, ν is an extremal Gibbs state iff $\forall f \in L^2(\nu), \lim_{t\to\infty} S^\nu(t)f = \nu(f)$ in $L^2(\nu)$, and different extremal Gibbs states are singular. These were key facts in proving the uniqueness of Gibbs states. Other relevant results can be found in [1, 2, 5].

(2) In [4], we also proved that every $\mu \in m_i(E)$ is a generalized convex combination of measures in $m_i^e(E)$.

Our next result is an application of theorem 2.1 and results concerning coupling. For coupling and fundamental theory, see [6, 10]. Here we only use some results given in [7, 12]. Denote by $\mathcal{S}$ the set of all time-invariant measurable subsets in Ω.

Theorem 2.2. *Let $\nu \in m_i(E)$*

(1) For $\mu \in m_1(E)$, $\lim_{t\to\infty} \mu T(t) = \nu$ strongly iff it holds uniformly. In particular, if the Markov process is transient, then the above convergence is also equivalent to $\lim_{t\to\infty} \mu S(t) = \nu$ uniformly.

(2) If $\nu \in m_i^e(E)$ and $\mu \ll \nu$, then $\lim_{t\to\infty} \mu T(t) = \nu$ *uniformly. In particular, if $m_i^e(E) = \{\nu\}$ with* $\mathrm{Supp}(\nu) = E$, *where* $\mathrm{Supp}(\nu)$ *is the support of ν, then* $\mathbf{S} = \{S(t),\ t \geqslant 0\}$ *is uniformly average ergodic.*

Proof.

(1) Obviously we only need to prove that strong convergence implies the uniform one. From the results in [7, 12], what we need to show is that $P_\mu|_{\mathcal{S}} = P_\nu|_{\mathcal{S}}$, where for $\pi \in m_1(E)$, $P_\pi = \int P_x \pi(\mathrm{d}x)$. Let $A \in \mathcal{S}$. Define $f_A(x) = P_x(A)$ for $x \in E$. Then by the shift invariance of A and the Markov property,

$$P_\mu(A) = \frac{1}{t}\int_0^t E^{P_\mu} P_{x_u}(A)\,\mathrm{d}u = \frac{1}{t}\int_0^t \mathrm{d}u \int S(u) f_A\,\mathrm{d}\mu = \mu T(t) f_A.$$

By the assumption, $\lim_{t\to\infty} \mu T(t) f_A = \nu(f_A) = P_\nu(A)$, proving the desired result.

(2) Let $\nu \in m_i^e(E)$ and $\mu \ll \nu$. For $A \in \mathcal{S}$, by theorem 2.1 (2) there is a sequence $t_n \to +\infty$, such that $T(t_n) f_A \to \nu(f_A) = P_\nu(A)$ $(n \to \infty)$ ν-a.s. Thus

$$\begin{aligned} |P_\mu(A) - P_\nu(A)| &= \left| \int T(t_n) f_A\,\mathrm{d}\mu - \nu(f_A) \right| \\ &\leqslant \int |T(t_n) f_A - \nu(f_A)|\,\mathrm{d}\mu \to 0 \qquad (n \to \infty) \end{aligned}$$

i.e. $P_\mu|_{\mathcal{S}} = P_\nu|_{\mathcal{S}}$. □

3. Complete convergence

Now we turn to the complete convergence for certain Markov processes. By complete convergence, here we mean that $\forall \mu \in m_1(E)$, $\mu S(t)$ converges to some $\nu \in m_i(E)$ weakly as $t \to \infty$. It is usually hard to prove such a type of convergence in general cases. Our results which follow suggest a method for studying this topic. To present them, we impose the following hypothesis:

(**H**) *For each $f \in C_b(E)$, each $x \in E$ and each $h > 0$*

$$\lim_{t\to\infty} |S(t+h) f(x) - S(t) f(x)| = 0.$$

A rough interpretation of this condition is that the difference between the laws of the system at any two different times with (arbitrary) fixed time difference will become negligible as the time tends to infinity, which is a property of shift invariance. As one can see from the proof of theorem 3.1, as a consequence of this condition, any limiting function of a subsequence of $S(t)f$ for $f \in C_b(E)$ is **S**-invariant (see theorem 3.1 for definition). We are now in a position to state our next main result.

Theorem 3.1. *Assume* (**H**) *and that* $\bigcup_{\nu\in m_i^e(E)} \mathrm{Supp}(\nu) = E$.

(1) If for every function f in a dense subset $C_{b,0}(E)$ of $C_b(E)$ and every sequence $t_n \to \infty$, there is an **S***-invariant function $f_0 \in C_b(E)$ (i.e. $S(t) f_0 = f_0$, $\forall t \geqslant 0$) and a subsequence t_{n_k}, such that*

$$\lim_{k\to\infty} S(t_{n_k}) f = f_0 \text{ holds on some dense } E_0 \subset E \tag{3.1}$$

then there is a $\pi \in m_1\big(m_i^e(E)\big)$, such that $\forall \mu \in m_1(E)$, $\lim_{t\to\infty} \mu S(t) = \int \nu \pi(\mathrm{d}\nu)$ *weakly. In particular, if $m_i^e(E) = \{\nu\}$ with* $\mathrm{Supp}(\nu) = E$, *then the process is ergodic.*

(2) If $\forall f \in C_{b,0}(E)$ and every sequence $t_n \to \infty$, the family of functions $\{S(t_n) f,\ n \geqslant 1\}$ on E is equi-continuous, then the same conclusion as above holds.

On the other hand, if for some different ν_1 *and* $\nu_2 \in m_i^e(E)$, $\mathrm{Supp}(\nu_1) \cap \mathrm{Supp}(\nu_2) \neq \emptyset$, *then complete convergence does not hold.*

Remark 3.1.

(i) A direct consequence of this theorem is that if a diffusion process on some bounded domain **D** has the Lebesgue measure as an invariant measure, then it is ergodic.
(ii) As we will show later, every non-explosive continuous-time Markov chain satisfies all the above conditions, except for $\cup_{\nu \in m_i^e(E)} \mathrm{Supp}(\nu) = E$. This is useful in further applications to IPS. Another situation in which the condition in (2) holds trivially is that when P_x is continuous in x in the uniform norm on $m_1(\Omega)$.

The following lemmas, which may have further applications, are key to proving the theorem.

Lemma 3.1. *Assume* (**H**). *Then for every* $\nu \in m_i^e(E)$ *and every bounded measurable function* f *on* E, $\lim_{t\to\infty} S(t)f = \nu(f)$ *in probability* ν.

Proof. Since $\nu \in m_i^e(E)$ and f is bounded, from therorem 2.1(2), $\forall \epsilon$ and $\delta > 0$, we can choose a sufficiently large T, such that

$$\left|\frac{1}{T}\int_0^T S(u)f\,\mathrm{d}u - \nu(f)\right| \mathrm{d}\nu < \frac{\epsilon\delta}{2}.$$

Note that

$$\nu(|S(t)f - \nu(f)| > \delta) \leqslant \Delta_1(t) + \Delta_2(t)$$

where

$$\Delta_1(t) = \nu\left(\left|\frac{1}{T}\int_0^T S(t+u)f\,\mathrm{d}u - \nu(f)\right| > \delta/2\right)$$

and

$$\Delta_2(t) = \nu\left(\left|\frac{1}{T}\int_0^T S(t+u)f\,\mathrm{d}u - S(t)\right| > \delta/2\right).$$

We have by the Markov property and the stationarity of ν that

$$\Delta_1(t) \leqslant \frac{2}{\delta}\int \left|\frac{1}{T}\int_0^T S(t+u)f\,\mathrm{d}u - \nu(f)\right| \mathrm{d}\nu \leqslant \frac{2}{\delta}\int \mathrm{d}\nu \left|\int_0^T S(u)f\,\mathrm{d}u - \nu(f)\right| < \epsilon$$

for large $T > 0$ and that, by (**H**), for such a fixed T,

$$\Delta_2(t) \leqslant \frac{2}{\delta}\int \mathrm{d}\nu \frac{1}{T}\int_0^T |S(t+u)f - S(t)f|\,\mathrm{d}u < \epsilon$$

for sufficiently large t, proving the lemma. □

Lemma 3.2. *If for each* $\nu \in m_i^e(E)$ *and each* $f \in C_b(E)$, $\lim_{t\to\infty} S(t)f = f_0$ ν*-a.s. for some* $f_0 \in C_b(E)$, *then for different* ν_1, $\nu_2 \in m_i^e(E)$, $\mathrm{Supp}(\nu_1) \cap \mathrm{Supp}(\nu_2) = \emptyset$.

Proof. Let $\nu \in m_i^e(E)$, f and $f_0 \in C_b(E)$ satisfy the above condition. Then f_0 is $\mathbf{S}^\nu$-invariant. Thus by theorem 2.1, $f_0 = \nu(f_0) = \nu(f)$ ν-a.s. By the continuity of f_0 we see that $f_0 = \nu(f)$ on $\mathrm{Supp}(\nu)$. The desired conclusion follows from this. □

Proof of theorem 3.1.

(1) Let $f,\ f_0,\ t_n$ and t_{nk} satisfy (3.1). Then f_0 is **S**-invariant by (**H**). From the proof of lemma 3.2 we see that $f_0 = \nu(f)$ on Supp(ν). This implies that $\lim_{t\to\infty} S(t)f(x) = \nu(f)$ on Supp(ν), i.e. $\lim_{t\to\infty} \delta_x S(t) = \nu$ for $x \in$ Supp(ν). The desired conclusion follows from this, lemma 3.2 and the assumption that $\cup_{\nu\in m_i^e(E)}\mathrm{Supp}(\nu) = E$.

(2) For any fixed $\nu \in m_i^e(E)$, by lemma 3.1, for each $f \in C_{b,0}(E)$ and every sequence $t_n \to \infty$, there is a subsequence t_{nk} such that $\lim_{k\to\infty} S(t_{n_k})f = \nu(f)$ ν-a.s. Now since $\{S(t_{n_k})f,\ k \geqslant 1\}$ is equi-continuous on E, it is easily seen that $\lim_{k\to\infty} S(t_{n_k})f = \nu(f)$ on Supp(ν). It then follows that $\lim_{t\to\infty} S(t)f = \nu(f)$ on Supp(ν). Then as for (1), the desired conclusion follows easily. The last conclusion is a direct consequence of lemma 3.2. □

As we stated at the beginning of the introduction, the study of our problems originates from understanding some ergodic behaviour of IPS. Duality is one of the powerful tools used to do this. Our next result is an application of theorem 3.1 concerning duality.

Corollary 3.1. *Let E_1 be a compact Polish space, E_2 be a Polish space,* $\mathbf{S}_i = \{S_i(t),\ t \geqslant 0\}$ *be a Markov process on E_i, $i = 1, 2$. We assume $\mathbf{S}_2$ satisfies the conditions of theorem 3.1. If for each f in some dense subset $C_{b,0}(E_1)$ of $C_b(E_1)$, there are a finite number of bounded measurable functions $f_{2,1}, \ldots, f_{2,l}$ on $E_1 \times E_2$ which are continuous in the second argument, such that for each $x_1 \in E_1$, we can find $x_{2,1}, \ldots, x_{2,l} \in E_2$ so that*

$$S_1(t)f(x_1) = \sum_{i=1}^{l} S_2(t)\ f_{2,i}(x_1, \cdot)(x_{2,i}) \qquad \forall t \geqslant 0$$

then the semigroup $\mathbf{S}_1$ is completely convergent.

Proof. By theorem 3.1, $\lim_{t\to\infty} S_2(t)f_{2,i}(x_1,\cdot)(x_{2,i})$ exists. Thus $\lim_{t\to\infty} \delta_{x_1} S_1(t)f = L_{x_1}(f)$ exists for any $x_1 \in E_1$ and $f \in C_{b,0}(E)$. Applying Riesz's representation theorem we see that there is a measure $\nu_{x_1} \in m_1(E_1)$ such that $L_{x_1}(f) = \nu_{x_1}(f)$. □

4. Applications to IPS

Now we sketch some applications of the above results to IPS. We mainly consider two classes of important IPS (spin flip and exclusion systems on Z^d) which are Feller–Markov processes $\Omega = D([0,\ \infty),\ E_1)$ with $E_1 = \{0,\ 1\}^{Z^d}$. It is expected that the following arguments can be applied to more general IPS. A spin flip system is characterized by a family of spin flip rates $\{c(i,\cdot),\ i \in Z^d\}$, where for each $i \in Z^d,\ c(i,\cdot)$ is a non-negative continuous function on E_1, which characterizes the change of state at site i. At any time, this change may take place at one site only. The generater A of such a system acts on any local function f on E as

$$\mathbf{A}f(\eta) = \sum_{i\in Z^d} c(i,\eta)[f(\eta^i) - f(\eta)]$$

where for $\eta \in E$ and $i \in Z^d$, $\eta^i \in E$ is defined by $\eta^i(j) = \eta(j)$ if $j \neq i$; $= 1-\eta(j)$ otherwise. For an exclusion system, however, at any time, only interchange of states between two sites may occur. The interchanges are characterized by a transition probability matrix $p(i,\ j)$ on Z^d. The generator of the system acts on a local function as

$$\mathbf{A}f(\eta) = \sum_{i,j\in Z^d} p(i,\ j)[f(\eta^{i,j}) - f(\eta)]$$

where for $\eta \in E$ and $i, j \in Z^d$, $\eta^{i,j} \in E$ is defined by $\eta^{i,j}(k) = \eta(j)$ if $k = i$; $= \eta(i)$ if $k = j$; $= \eta(k)$ otherwise. For precise description of these systems see ([9], ch III and VIII).

As for the first application, we present a result concerning extremality of a class of invariant measures of the systems.

Theorem 4.1.

(1) For a spin flip system with strictly positive spin flip rates $c(\cdot,\cdot)$, let ν be an invariant measure of it. Then ν is extremal iff every ν-integrable function f on E that satisfies

$$f(\eta^i) = f(\eta) \qquad \forall i \ \nu\text{-a.s.} \tag{4.1}$$

is ν-a.s. constant.

(2) For an exclusion process with irreducible $p(i, j)$, an invariant measure ν of it is extremal iff every ν-integrable function f that satisfies

$$f(\eta^{i,j}) = f(\eta) \qquad \forall i, j \ \nu\text{-a.s.}$$

is ν-a.s. constant.

In particular, if an invariant measure ν of one of the above systems satisfies a zero-one law on the tail σ-algebra $\mathcal{T}$ on E, i.e. $\nu(A) = 0$ or $\forall A \in \mathcal{T}$, then ν is extremal.

Sketch of proof. We only prove the first conclusion. The proof of the second conclusion is similar. For sufficiency, let f be a bounded and $\mathbf{S}^\nu$-invariant function on E. Then an argument similar to the one used in the proof of lemma IV.4.3 shows that

$$0 = \int f\mathbf{A}f \, \mathrm{d}\nu = -\frac{1}{2}\sum_{i\in Z^d}\int c(i,\eta)[f(\eta^i) - f(\eta)]^2 \, \mathrm{d}\nu.$$

Thus, since $c(\cdot,\cdot)$ is strictly positive, $f(\eta^i) = f(\eta)$ $\forall i$ ν-a.s. By assumptions, $f =$ constant ν-a.s. on E. From theorem 2.1 we see that ν is extremal.

For necessity, let ν be an extremal invariant measure of the system, f be a ν-integrable function on E satisfying (4.1). By a standard argument, we can assume that f is bounded and satisfies

$$\alpha \equiv \int f \, \mathrm{d}\nu > 0.$$

Define ν_1 by

$$\mathrm{d}\nu_1 = \alpha^{-1} f \, \mathrm{d}\nu.$$

It is not hard to check that $\forall f \in \mathcal{D}(\mathbf{A})$—the domain of $\mathbf{A}$

$$\int \mathbf{A}f \, \mathrm{d}\nu_1 = 0$$

i.e. $\nu_1 \in m_i(E)$. Since $\nu_1 \ll \nu$, by theorem 2.1 we obtain that $\nu_1 = \nu$. It follows that $f =$ constant ν-a.s. on E.

The last conclusion of the theorem follows from the fact that a function f satisfying (4.1) is $\mathcal{T}$-measurable and the assumption on ν.

Remark 4.1. A simple consequence of the theorem is that if a product measure ν on E with constant marginal density is invariant for the system under consideration, then it is extremal.

Now we turn to applications of the results presented in sections 2 and 3 to convergence to invariant measures of the systems. To do this, we consider only those spin flip systems with spin flip rates given by (4.3)–(4.10) in ([9], ch III, section 4) and symmetric exclusion systems. Most typical and interesting systems, such as contact, voter and anti-voter models, etc, are included. One of the fundamental facts we will use for these systems is that if we denote by $\mathbf{S}_1$ the Markov semigroup of such a system, then there is a continuous time non-explosive Markov chain $\mathbf{S}_2$ on some countable state space E_2, and a non-negative function V on E_2, such that for each local function f on E_1 (i.e. for some finite $\Lambda \subset Z^d$, $f(\eta)$ depends only on $\{\eta(i), i \in \Lambda\}$), there are bounded measurable functions $f_{2,i}$, $1 \leqslant i \leqslant l$, on $E_1 \times E_2$, such that for each $\eta \in E_1$, there are $y_1, \dots, y_l \in E_2$ so that

$$S_1(t)\, f(\eta) = \sum_{i=1}^{l} E^{y_i} f_{2,i}\,(\eta, y_t) \exp\left\{-\int_0^t V(y_u)\,\mathrm{d}u\right\} \qquad \forall t \geqslant 0.$$

Note that the semigroup

$$S^V(t)g(y) = E^y g(y_t) \exp\left\{-\int_0^t V(y_u)\,\mathrm{d}u\right\}$$

on E_2 is sub-Markov, i.e. $S^V(t)1 \leqslant 1$, we can add a point Δ to E_2 to make $\mathbf{S}^V$ a nonexplosive Markov chain on E_2 (cf [6]). Thus we see that (3.1) is satisfied. Therefore if we can prove the complete convergence for the dual Markov chain, then we will know that the corresponding IPS is completely convergent. The following proposition gives a useful result for Markov chains.

Proposition . *Let* $\mathbf{S} = \{S(t), t \geqslant 0\}$ *be a nonexplosive Markov chain on a countable state space E. Then* (**H**) *is satisfied.*

Proof. Let $f \in C_b(E)$ and $h_0 > 0$. Consider the discrete time Markov chain $\{S(nh_0), n \geqslant 1\}$. From [11] we know that

$$\lim_{n\to\infty} \sup_x |S(nh_0 + h_0) f(x) - S(nh_0) f(x)| = 0.$$

Define $\phi(t) = |S(t+h) f(x) - S(t) f(x)|$. Then $\phi \in C_b([0, \infty))$, and $\forall h > h_0$

$$\lim_{n\to\infty} \phi(nh) \leqslant \lim_{n\to\infty} \sup_x |S(nh_0 + h_0) f(x) - S(nh_0) f(x)| = 0.$$

Thus from the result in [8] we see that $\lim_{t\to\infty} \phi(t) = 0$. □

If the state space E of the Markov chain is equipped with the discrete topology, then obviously the family $\{S(t)f, t \geqslant 0\}$ is equi-continuous on E for each $f \in C_b(E)$. Therefore the only condition we need to check for complete convergence is that $\cup_{\nu \in m_i^e(E)} \mathrm{Supp}(\nu) = E$. In particular, if E can be decomposed into several subsets $E_1, \dots, E_r$, such that either E_i consists of absorbing states, or the chain is ergodic on E_i, then the chain is completely convergent; see the following example for a discussion.

Example 1. For simplicity, consider a one-dimensional linear voter model, which is a spin flip system on Z with spin flip rates given by

$$c(i, \eta) = \tfrac{1}{2}|\{j = -1, 1 : \eta(i+j) \neq \eta(i)\}|$$

where $|A|$ denotes the cardinality of $A \subset Z$. Then its dual Markov chain is the one with state space $\mathbf{Y}$ = the collection of all finite subsets of Z, and with transition rates given by

$$q(A, B) = \begin{cases} \frac{1}{2} & \text{if } A \neq \emptyset \text{ and } B = (A - x) \cup \{x-1\} \text{ or } (A - x) \cup \{x+1\} \\ |A| & \text{if } A \neq \emptyset \text{ and } B = A \\ 0 & \text{otherwise} \end{cases} \quad .$$

It is not hard to check that $\emptyset$ and ∞ are absorbing states of this chain, and the chain is irreducible on $Y_0 = Y - \{\emptyset, \{\infty\}\}$. Using an argument similar to the one appearing in the proof of theorem 4.58 in [6] we can see that this Markov chain on Y_0 is ergodic. Combining this with the above proposition one can easily show that this Markov chain, and hence the voter model, is completely convergent. Extention to more general voter models is direct.

A similar argument applies to linear anti-voter models. The next is an example for which complete convergence does not hold.

Example 2. Now consider an exclusion system with the states interchanging rate between i and $j \in Z^d$ given by $p(i, j)$. It is known that if $p(\cdot, \cdot)$ is doubly stochastic, i.e. $\sum_j p(i, j) = \sum_i p(i, j) = 1\ \forall i,\ j \in Z^d$, then every product measure ν_α with constant density α is invariant for this system. Thus by theorem 3.1, it cannot have complete convergence.

The following is a direct consequence of the above proposition.

Corollary 4.1. *For every spin flip or exclusion system described above,* (**H**) *holds.*

Combining this with lemma 3.1 we have the following interesting application which is relevant to the open problem presented in [9]:

Corollary 4.2. *Given a spin flip or exclusion system described as above with state space* $E = \{0, 1\}^{Z^d}$. *For* $\mu \in m_1(E)$ *and sequence* $t_n \to \infty$, *if* $\lim_{n\to\infty} \mu S(t_n)$ *exists in the weak topology, then the limit is in* $m_i(E)$.

Proof. The assertion follows from (**H**) easily. □

Acknowledgments

We are grateful to the referee for helpful comments and suggestions. Supported by the NSF of China.

References

[1] Albeverio S, Bogachev V and Rochner M 1999 On uniquesess of invariant measures for finite- and infinite-dimensional diffusions *Comment. Pure Appl. Math.* **LII** 325

[2] Albeverio S, Bogachev V and Rochner M 1997 Ergodicity of semigroups and extremality of Gibbs states *J. Funct. Anal.* **144** 394

[3] Andjel E 1990 Ergodic and mixing properties of equilibrium measures for Markov processes *Trans. Am. Math. Soc.* **318** 601

[4] Chen J W Shift ergodicity for stationary distributions of Markov processes *Preprint*

[5] Chen J W 1999 A variational principle for Markov processes *J. Stat. Phys.* **96** 1359

[6] Chen M F 1992 *From Markov Chains to Non-Equilibrium Particle Systems* (Singapore: World Scientific)

[7] Cranston M and Greren A 1995 Coupling and harmonic functions in the case of continuous time Markov processes *Stoch. Processes Appl.* **60** 261

[8] Kingman J F C 1963 Ergodic properties of continuous-time Markov processes and their discrete time skeleton *Proc. Lond. Math. Soc.* **13** 593

[9] Liggett T M 1985 *Interacting Particle Systems* (New York: Springer)

[10] Lindval T 1992 *Lectures on the Coupling Method* (New York: Wiley)

[11] Revuz D 1984 *Markov Chains* (Amsterdam: North-Holland)

[12] Therisson H 1994 Shift-coupling in continuous-time *Probab. Theor. Relat. Fields.* **99** 477