PROCEEDINGS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

On convergence of chains with occupational self–interactions

Pierre Del Moral and Laurent Miclo

Proc. R. Soc. Lond. A 2004 **460**, doi: 10.1098/rspa.2003.1245, published 8 January 2004

10.1098/rspa.2003.1245

On convergence of chains with occupational self-interactions

By Pierre Del Moral and Laurent Miclo

Laboratoire de Statistique et Probabilités, UMR 5583, CNRS et Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4, France
(delmoral@math.ups-tlse.fr; miclo@math.ups-tlse.fr)

Received 17 February 2003; accepted 1 September 2003; published online 14 November 2003

We consider stochastic chains on abstract measurable spaces whose evolution at any given time depends on the present position and on the occupation measure created by the path up to this instant. This generalization of reinforced random walks enables us to impose conditions ensuring $\mathbb{L}^p$, $p \geqslant 1$, or a.s. convergence of the empirical measures towards some fixed point of a probability-valued dynamical system. We present two sets of hypotheses based on weak contraction properties, leading to two different proofs, but in both situations the rates of convergence are optimal in the examined level of generality.

Keywords: reinforced walk; weak contraction; $\mathbb{L}^p$-rate of convergence; a.s. convergence; Poisson equation; resolvent estimate

1. Introduction

In a recent article (Del Moral & Miclo 2002), we have considered stochastic chains whose time evolution depends on the empirical measure created by the past trajectory and also more specifically on the last state visited. Our objective in introducing such algorithms was to 'numerically' approximate fixed points of certain dynamical systems taking values in spaces of probability measures. We refer the reader to this former paper for more information about our initial motivations and for examples coming from genetical algorithms or nonlinear filtering theory. Our goal here is to improve qualitatively and quantitatively our previous results. In particular, we will prove almost-sure (a.s.) asymptotic behaviours.

(*a*) *The framework*

As we will work in an abstract setting, let us introduce some general notation in order to describe the considered model. On a measurable space $(E, \mathcal{E})$, we consider $\mathcal{B}(E)$, $\mathcal{P}(E)$ and $\mathcal{K}(E)$, the set of numerical bounded and measurable functions, of probability measures and of probability kernels, respectively. We recall that the latter are mappings K from $E \times \mathcal{E}$ to $[0, 1]$ verifying the following two properties:

(i) for any $x \in E$, $\mathcal{E} \ni A \mapsto K(x, A)$ belongs to $\mathcal{P}(E)$;

(ii) for any $A \in \mathcal{E}$, $E \ni x \mapsto K(x, A)$ belongs to $\mathcal{B}(E)$.

As usual, such an element $K \in \mathcal{K}(E)$ can be seen as a right-acting (respectively, left-acting) mapping on $\mathcal{B}(E)$ (respectively, $\mathcal{P}(E)$) by

$$\forall f \in \mathcal{B}(E),\ \forall x \in E, \quad K[f](x) := \int f(y) K(x, \mathrm{d}y)$$

(respectively, $\forall p \in \mathcal{B}(E)$, $\forall A \in \mathcal{E}$, $pK[A] := \int K(x, A) p(\mathrm{d}x)$).

Our main object of interest is a mapping $\mathbb{K} : \mathcal{P}(E) \ni m \mapsto K_m \in \mathcal{K}(E)$ and our original motivation in Del Moral & Miclo (2002) was to find a $\Pi \in \mathcal{P}(E)$ such that $\Pi K_\Pi = \Pi$. More precisely, our objective was to derive a new algorithm which permits us to find such an invariant probability Π, or at least to give conditions on $\mathbb{K}$ ensuring the existence and uniqueness of this solution Π and next the convergence in some sense of our algorithm. Heuristically, it was based on a stochastic chain $X := (X_n)_{n \geqslant 0}$ taking values in E, defined in some probability space $(\Omega, \mathcal{T}, \mathbb{P})$ and whose evolution is given by the conditional probabilities

$$\forall n \in \mathbb{N}^*,\ \forall A \in \mathcal{E}, \quad \mathbb{P}[X_n \in A \mid X_0, X_1, \ldots, X_{n-1}] = K_{S_{n-1}}(X_{n-1}, A) \tag{1.1}$$

($\mathbb{P}$-a.s.), where, for all $n \in \mathbb{N}$, we have defined

$$S_n = \frac{1}{n+1} \sum_{0 \leqslant i \leqslant n} \delta_{X_i} \in \mathcal{P}(E).$$

Indeed, there is no problem in constructing such a chain, as soon as, for all $n \in \mathbb{N}$ and all $A \in \mathcal{E}$, the mapping

$$E^{n+1} \ni (x_0, \ldots, x_n) \mapsto K_{(1/(n+1)) \sum_{0 \leqslant i \leqslant n} \delta_{x_i}}(x_n, A) \tag{1.2}$$

is $\mathcal{E}^{\otimes(n+1)}$-measurable and then if the law of X_0 is given, the law of X is uniquely determined (on $(E^{\mathbb{N}}, \mathcal{E}^{\otimes \mathbb{N}})$). So implicitly, we will always assume this measurability property of $\mathbb{K}$, to be sure to have at our disposal a Markov chain $(X_n, S_n)_{n \in \mathbb{N}}$ whose evolution is defined by (1.1). But clearly, much stronger assumptions on $\mathbb{K}$ are needed to ensure that our algorithm is working, namely that S_n converges in probability to the expected Π for n large.

In order to make this convergence precise, let us call a subset $\mathcal{F} \subset \mathcal{B}(E)$ a test-functions collection if for all $f \in \mathcal{F}$ its supremum norm satisfies $\|f\|_\infty \leqslant 1$ and if $d_{\mathcal{F}}$ is a complete metric on $\mathcal{P}(E)$, where by definition

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

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

$$\mathcal{F} = \{f \in \mathcal{B}(E) : \|f\|_\infty \leqslant 1\}.$$

In this case the distance $d_{\mathcal{F}}$ is actually complete, since it is given by the total variation norm.

One can also recover the latter norm by considering, for instance,

$$\mathcal{F} = \{f \in C_{\mathrm{b}}(E) : \|f\|_\infty \leqslant 1\},$$

where $C_{\mathrm{b}}(E)$ 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 $(M_n)_{n\in\mathbb{N}}$ is a sequence of real numbers larger than unity and diverging to $+\infty$ and $(f_n)_{n\in\mathbb{N}}$ is a weak convergence determining sequence of $C_{\mathrm{b}}(E)\setminus\{0\}$ (recall that this means that a sequence of probability measures $(p_l)_{l\in\mathbb{N}}$ is weakly convergent if and only if for all fixed $n \in \mathbb{N}$, $(p_l[f_n])_{l\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$).

If $\mathcal{F}$ is a test-functions collection, we will always endow it with the σ-algebra $\mathcal{E}(\mathcal{F})$ naturally generated by the mappings

$$\mathcal{F} \ni f \mapsto p[f] \in \mathbb{R}$$

for all $p \in \mathcal{P}(E)$.

In a very reciprocal way, we endow $\mathcal{P}(E)$ with the weak topology associated to its duality with $\mathcal{F}$ (i.e. a basis of neighbourhoods of a probability measure $p \in \mathcal{P}(E)$ is given by the sets $\{q \in \mathcal{P}(E) : \forall 0 \leqslant i \leqslant n, |q(f_i) - p(f_i)| < \epsilon\}$, where $n \in \mathbb{N}$, $f_1,\dots,f_n \in \mathcal{F}$ and $\epsilon > 0$).

From now on, we assume that we are given a test-functions collection $\mathcal{F}$ and we will impose convenient regularity conditions on $\mathbb{K}$ with respect to $\mathcal{F}$ to obtain the expected convergence in probability (which is equivalent to the convergence in probability of $S_n[f]$ to $\Pi[f]$, for every fixed $f \in \mathcal{F}$) or stronger asymptotic behaviours.

(*b*) *Plan of the paper*

In next section we will give a simple condition ensuring an $\mathbb{L}^2$ convergence by a direct proof. This method could be extended to obtain more general $\mathbb{L}^p$, $p \geqslant 1$, or a.s. convergence in the same way as it is done in § 4, but keeping the arguments of this section.† Nevertheless, we prefer to present a little more involved approach, which enables us to extend the hypotheses, at least if one has at his disposal more *a priori* information on $\mathbb{K}$ and its associated invariant probabilities. So, in § 3, we deal with some prerequisites on resolvent solutions of Poisson's equation and in § 4 we develop a second proof to obtain the required convergences. These results improve the ones derived in our previous article (Del Moral & Miclo 2002), in particular, at the level of obtained rates of convergence (even in the restrictive $\mathbb{L}^2$ sense, which was the only convergence treated in our previous article). A theoretical example is next considered in order to illustrate the advantage of our second set of hypotheses and other features of our conditions. We add two appendices, one concerning the discrete version of a traditional differential inequality which is used throughout the paper, and the other discussing further relations between our hypotheses.

2. A direct $\mathbb{L}^2$-approach

Our objective here is to obtain by a simple proof $\mathbb{L}^2$ bounds of convergence for the algorithm presented in § 1 under appropriate conditions on $\mathbb{K}$. More precisely, we will assume in this section that the following hypothesis is fulfilled.

† For more details, contact the authors to obtain the corresponding LaTeX file.

Hypothesis 2.1. *For any $(f,m) \in \mathcal{F} \times \mathcal{P}(E)$ there exist two non-negative measures $\mu_{f,m}$ and $\nu_{f,m}$ on the measurable space $(\mathcal{F}, \mathcal{E}(\mathcal{F}))$ satisfying*

$$\lambda_1 := \sup\{\mu_{f,m}(\mathbb{1}) : f \in \mathcal{F}, m \in \mathcal{P}(E)\} < 1,$$
$$\lambda_2 := \sup\{\nu_{f,m}(\mathbb{1}) : f \in \mathcal{F}, m \in \mathcal{P}(E)\} < 1 - \lambda_1,$$

and such that for any $(f,m) \in \mathcal{F} \times \mathcal{P}(E)$ we are assured of the following inequalities for all $p_1, p_2 \in \mathcal{P}(E)$,

$$|p_1 K_m[f] - p_2 K_m[f]| \leqslant \int |p_1[h] - p_2[h]| \mu_{f,m}(\mathrm{d}h),$$
$$|mK_{p_1}[f] - mK_{p_2}[f]| \leqslant \int |p_1[h] - p_2[h]| \nu_{f,p_1}(\mathrm{d}h).$$

We note that, to get the latter upper bound, it is sufficient to consider Dirac masses for m and this second condition can equivalently be written as

$$\forall f \in \mathcal{F},\ \forall p_1, p_2 \in \mathcal{P}(E), \quad \|(K_{p_1} - K_{p_2})[f]\|_\infty \leqslant \int |p_1[h] - p_2[h]| \nu_{f,p_1}(\mathrm{d}h).$$

The interpretation of the above inequalities is quite obvious, ν serves to quantify a contraction assumption on $\mathbb{K}$ as a mapping from $\mathcal{P}(E)$ to $\mathcal{K}(E)$, while μ is related to mixing properties of its kernels K_m, for $m \in \mathcal{P}(E)$.

We also remark that under some supplementary mild conditions, the measurability of the mappings presented in (1.2) is a consequence of hypothesis 2.1. For instance, this is the case if we assume that $\mathcal{F}$ is at most denumerable, that $\mathcal{E}$ is the σ-field generated by $\mathcal{F}$ and that the $\|\cdot\|_\infty$-closure of the vector space generated by $\mathcal{F}$ is an algebra. To show this affirmation, we introduce a metric d on E by

$$\forall x, y \in E, \quad d(x,y) := d_{\mathcal{F}}(\delta_x, \delta_y) = \sup_{f \in \mathcal{F}} |f(x) - f(y)|.$$

Note that then $\mathcal{E}$ is nothing but the corresponding Borelian σ-field. With respect to the induced product topology and due to the above inequalities, it appears easily that the mappings defined in (1.2) are continuous, if we replace A by a function $f \in \mathcal{F}$. A monotonous class theorem then enables us to conclude.

Our main result in this section can now be stated as follows.

Theorem 2.2. *Under hypothesis 2.1 on $\mathcal{F}$ and $\mathbb{K}$, there exists a unique fixed point $\Pi \in \mathcal{P}(E)$ for the equation $\Pi = \Pi K_\Pi$ and the occupation measures S_n constructed in § 1 weakly converge to Π in probability for large n. Furthermore, if we define $\Lambda = \lambda_2/(1-\lambda_1) \in [0,1)$, we have three different types of upper bounds for the $\mathbb{L}_2$-mean-error decays: there exists a constant $c \geqslant 0$ (depending only on Λ) such that, for any $f \in \mathcal{F}$ and any n larger than two,*

$$\Lambda < \tfrac{1}{2} \Rightarrow \mathbb{E}((S_n[f] - \Pi[f])^2) \leqslant \frac{c}{n},$$
$$\Lambda = \tfrac{1}{2} \Rightarrow \mathbb{E}((S_n[f] - \Pi[f])^2) \leqslant \frac{c\ln(n)}{n},$$
$$\Lambda > \tfrac{1}{2} \Rightarrow \mathbb{E}((S_n[f] - \Pi[f])^2) \leqslant \frac{c}{n^{2(1-\Lambda)}}.$$

As is shown by an example given in Del Moral & Miclo (2002), these rates can be optimal in certain situations.

The proof of this result will consist of two steps, corresponding to its first and last assertions, one deterministic and the other stochastic.

(*a*) *Existence and uniqueness of Π*

We begin by noting that, for any $m \in \mathcal{P}(E)$, there exists a unique invariant probability measure $\varpi(m)$ relative to the kernel K_m. Indeed, taking suprema in the first inequality of hypothesis 2.1, we obtain

$$\forall p_1, p_2 \in \mathcal{P}(E), \quad d_{\mathcal{F}}(p_1 K_m, p_2 K_m) \leqslant \lambda_1 d_{\mathcal{F}}(p_1, p_2), \tag{2.1}$$

so that a usual fixed point theorem gives us the expected result.

In the same way, to show the uniqueness and existence of Π, it is sufficient to verify that

$$\forall m_1, m_2 \in \mathcal{P}(E), \quad d_{\mathcal{F}}(\varpi(m_1), \varpi(m_2)) \leqslant \frac{\lambda_1}{1-\lambda_2} d_{\mathcal{F}}(m_1, m_2), \tag{2.2}$$

since there then exists a unique fixed point Π such that $\Pi = \varpi(\Pi)$ and this equation is in fact easily seen to be equivalent to $\Pi = \Pi K_{\Pi}$.

To prove the above bound, we come back to (2.1), which shows by iteration that, for all $n \in \mathbb{N}$ and all $m, p_1, p_2 \in \mathcal{P}(E)$,

$$d_{\mathcal{F}}(p_1 K_m^n, p_2 K_m^n) \leqslant \lambda_1^n d_{\mathcal{F}}(p_1, p_2)$$

(the convention $K_m^0 = \mathrm{Id}$ will always be enforced), so that, for all $n \in \mathbb{N}$ and all $m_1, m_2 \in \mathcal{P}(E)$,

$$\begin{aligned}
d_{\mathcal{F}}(\varpi(m_1), \varpi(m_2)) &\leqslant d_{\mathcal{F}}(\varpi(m_1), \varpi(m_1) K_{m_2}^n) + d_{\mathcal{F}}(\varpi(m_1) K_{m_2}^n, \varpi(m_2)) \\
&\leqslant \sum_{0 \leqslant k \leqslant n-1} d_{\mathcal{F}}(\varpi(m_1) K_{m_1}^{n-k} K_{m_2}^k, \varpi(m_1) K_{m_1}^{n-k-1} K_{m_2}^{k+1}) \\
&\quad + d_{\mathcal{F}}(\varpi(m_1) K_{m_2}^n, \varpi(m_2)) \\
&= \sum_{0 \leqslant k \leqslant n-1} d_{\mathcal{F}}(\varpi(m_1) K_{m_1} K_{m_2}^k, \varpi(m_1) K_{m_2} K_{m_2}^k) \\
&\quad + d_{\mathcal{F}}(\varpi(m_1) K_{m_2}^n, \varpi(m_2) K_{m_2}^n) \\
&\leqslant \sum_{0 \leqslant k \leqslant n-1} \lambda_1^k d_{\mathcal{F}}(\varpi(m_1) K_{m_1}, \varpi(m_1) K_{m_2}) \\
&\quad + \lambda_1^n d_{\mathcal{F}}(\varpi(m_1), \varpi(m_2)) \\
&\leqslant \sum_{0 \leqslant k \leqslant n-1} \lambda_1^k \lambda_2 d_{\mathcal{F}}(m_1, m_2) + \lambda_1^n d_{\mathcal{F}}(\varpi(m_1), \varpi(m_2)),
\end{aligned}$$

where we have used the relation

$$\forall m, p_1, p_2 \in \mathcal{P}(E), \quad d_{\mathcal{F}}(m K_{p_1}, m K_{p_2}) \leqslant \lambda_2 d_{\mathcal{F}}(p_1, p_2),$$

which is deduced from the second inequality of hypothesis 2.1, also by considering suprema over $\mathcal{F}$. Now letting n tend to infinity, we get (2.2).

(b) $\mathbb{L}^2$-*estimates of convergence*

Our proof of the bounds presented in theorem 2.2 is based on a recurrence relation between some quantities, which we now introduce. For all $n \in \mathbb{N}$ and $f, g \in \mathcal{F}$, we note

$$I_n(f) := (n+1)^2 \mathbb{E}[(S_n[f] - \Pi[f])^2],$$
$$J_n(f,g) := n\mathbb{E}[(S_{n-1}[f] - \Pi[f])(g(X_n) - \Pi[g])]$$

(the latter term is equal to zero for $n = 0$), and we consider next

$$I_n := \sup_{f \in \mathcal{F}} I_n(f),$$
$$J_n := \sup_{f,g \in \mathcal{F}} |J_n(f,g)|.$$

Our estimates will be deduced from the following system.

Lemma 2.3. *For all $n \in \mathbb{N}$, we have*

$$I_{n+1} \leqslant \left(1 + \frac{2\lambda_2}{n+1}\right) I_n + 2\lambda_1 J_n + 4(1 + \lambda_1),$$
$$J_{n+1} \leqslant \frac{\lambda_2}{n+1} I_n + \lambda_1 J_n + 4\lambda_1.$$

Proof. Let us start with a given function $f \in \mathcal{F}$. For any $n \in \mathbb{N}$, we have

$$\begin{aligned} I_{n+1}(f) &= \mathbb{E}\left[\left(\sum_{0 \leqslant i \leqslant n+1} \{f(X_i) - \Pi(f)\}\right)^2\right] \\ &= I_n(f) + 2J_{n+1}(f,f) + \mathbb{E}[(f(X_{n+1}) - \Pi(f))^2] \\ &\leqslant I_n(f) + 2J_{n+1}(f,f) + 4. \end{aligned}$$

But we also have for any $n \in \mathbb{N}$ and $f, g \in \mathcal{F}$, using (1.1),

$$\begin{aligned} J_{n+1}(f,g) &= (n+1)\mathbb{E}[(S_n[f] - \Pi[f])(K_{S_n}[g](X_n) - \Pi[g])] \\ &= (n+1)\mathbb{E}[(S_n[f] - \Pi[f])(K_{S_n}[g](X_n) - K_\Pi[g](X_n))] \\ &\quad + (n+1)\mathbb{E}[(S_n[f] - \Pi[f])(K_\Pi[g](X_n) - \Pi[g])]. \end{aligned}$$

We evaluate each of the last two terms separately. For the first term, its absolute value is bounded above by

$$\begin{aligned} &(n+1)\mathbb{E}\left[|S_n[f] - \Pi[f]| \int |S_n[h] - \Pi[h]| \nu_{g,\Pi}(\mathrm{d}h)\right] \\ &\quad = (n+1) \int \mathbb{E}[|S_n[f] - \Pi[f]||S_n[h] - \Pi[h]|] \nu_{g,\Pi}(\mathrm{d}h) \\ &\quad \leqslant \frac{1}{n+1} \int \sqrt{I_n(f) I_n(h)} \nu_{g,\Pi}(\mathrm{d}h) \\ &\quad \leqslant \lambda_2 \frac{I_n}{n+1}. \end{aligned}$$

To treat the second term, note that the first bound of hypothesis 2.1 can be rewritten as

$$\forall (f,m) \in \mathcal{F} \times \mathcal{P}(E), \quad |qK_m[f]| \leqslant \int |q[h]| \mu_{f,m}(\mathrm{d}h),$$

where q is an arbitrary measure on $(E, \mathcal{E})$ such that $q(\mathbb{1}) = 0$.

Applying this to $m = \Pi$ and to the measure q_n defined by

$$\forall A \in \mathcal{E}, \quad q_n(A) = (n+1)\mathbb{E}[(S_n[f] - \Pi[f])(\mathbb{1}_A - \Pi(A))],$$

clearly verifying $q_n(\mathbb{1}) = 0$, we obtain

$$\begin{aligned}
(n+1)|\mathbb{E}[(S_n[f] - \Pi[f])&(K_\Pi[g](X_n) - \Pi[g])]| \\
&= |q_n[K_\Pi(g)]| \\
&\leqslant \int |q_n(h)| \mu_{g,\Pi}(\mathrm{d}h) \\
&= (n+1) \int |\mathbb{E}[(S_n[f] - \Pi[f])(h(X_n) - \Pi[h])]| \mu_{g,\Pi}(\mathrm{d}h) \\
&= \int |\mathbb{E}[\{n(S_{n-1}[f] - \Pi[f]) + f(X_n) - \Pi[f]\}(h(X_n) - \Pi[h])]| \mu_{g,\Pi}(\mathrm{d}h) \\
&\leqslant \int 4 + |J_n(f,h)| \mu_{g,\Pi}(\mathrm{d}h) \\
&\leqslant \lambda_1(4 + J_n)
\end{aligned}$$

and the above lemma follows at once. ■

It seems that the simplest way to study the asymptotic behaviour of this sequence $(I_n, J_n)_{n\in\mathbb{N}}$ is to introduce

$$\forall n \in \mathbb{N}, \quad L_n := I_n + \frac{2\lambda_1}{1-\lambda_1} J_n,$$

because we are assured of

$$\begin{aligned}
L_{n+1} &\leqslant \left(1 + \frac{2\lambda_2}{n+1}\right) I_n + 2\lambda_1 J_n + 4(1+\lambda_1) \\
&\quad + \frac{2\lambda_1\lambda_2}{(1-\lambda_1)(n+1)} I_n + \frac{2\lambda_1^2}{1-\lambda_1} J_n + \frac{4\lambda_1^2}{1-\lambda_1} \\
&= \left(1 + \frac{2\Lambda}{n+1}\right) I_n + \frac{2\lambda_1}{1-\lambda_1} J_n + \frac{4}{1-\lambda_1} \\
&\leqslant \left(1 + \frac{2\Lambda}{n+1}\right) L_n + \frac{4}{1-\lambda_1}.
\end{aligned}$$

We are therefore naturally led to the first technical lemma of Appendix A (with $\alpha = 2\Lambda$ and $\beta = 0$) to see there exists a constant $c > 0$ such that, for n large enough, we have

$$\begin{aligned}
\Lambda < \tfrac{1}{2} &\Rightarrow L_n \leqslant cn, \\
\Lambda = \tfrac{1}{2} &\Rightarrow L_n \leqslant cn\ln(n), \\
\Lambda > \tfrac{1}{2} &\Rightarrow L_n \leqslant cn^{2\Lambda}.
\end{aligned}$$

In particular, these bounds are also valid for I_n, which is just the result announced in theorem 2.2, once we take into account the included factor $(n+1)^2$.

3. A resolvent method

In the classical situation of a mixing probability kernel, the simplest way to derive the law of large numbers, or the central limit theorem, for the occupation measure, is certainly to use a resolvent method. Our objective here is to adapt it to our setting and to see what we obtain as an ersatz.

To do so we need to modify our set of hypotheses, and the first condition we impose is as follows.

Hypothesis 3.1. *There exist $L_0 \geqslant 0$, $n_0 \in \mathbb{N}^*$ and $0 \leqslant \lambda < 1$ such that, for all $p, m_1, m_2 \in \mathcal{P}(E)$, we are assured of*

$$d_{\mathcal{F}}(m_1 K_p, m_2 K_p) \leqslant L_0 d_{\mathcal{F}}(m_1, m_2),$$
$$d_{\mathcal{F}}(m_1 K_p^{n_0}, m_2 K_p^{n_0}) \leqslant \lambda d_{\mathcal{F}}(m_1, m_2).$$

The first bound will be necessary to get round the lack of assumed stability of $\mathcal{F}$ by the kernels of $\mathbb{K}$. As in § 2 *a*, it readily follows from the second bound that, for all $p \in \mathcal{P}(E)$, there exists a unique invariant probability $\varpi(p)$ for the kernel $K_p^{n_0}$. This measure $\varpi(p)$ is also invariant for K_p. Since we have

$$(\varpi(p)K_p)K_p^{n_0} = (\varpi(p)K_p^{n_0})K_p = \varpi(p)K_p,$$

$\varpi(p)K_p$ is an invariant probability for $K_p^{n_0}$, and thus by uniqueness $\varpi(p)K_p = \varpi(p)$. Clearly, $\varpi(p)$ is indeed the unique invariant probability for K_p.

In order to present our second hypothesis, let us leave $\mathcal{P}(E)$ and denote by $\mathcal{M}_0(E)$ the vector space of signed bounded measures m on $(E, \mathcal{E})$ with zero total mass. We extend the distance $d_{\mathcal{F}}$ by considering on $\mathcal{M}_0(E)$ the norm $\|\cdot\|_{\mathcal{F}}$ defined by

$$\forall m \in \mathcal{M}_0(E), \quad \|m\|_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |m[f]|.$$

We will also have resort to the total variation norm, which corresponds to the special case $\mathcal{F} = \{f \in \mathcal{B}(E) : \|f\|_\infty \leqslant 1\}$,

$$\forall m \in \mathcal{M}_0(E), \quad \|m\|_{\mathrm{tv}} := \sup_{\substack{f \in \mathcal{B}(E), \\ \|f\|_\infty \leqslant 1}} |m[f]|.$$

We will furthermore assume that the following hypothesis holds.

Hypothesis 3.2. *There exist two constants $L_1, L_2 \geqslant 0$ such that, for all $m \in \mathcal{M}_0(E)$ and all $p_1, p_2 \in \mathcal{P}(E)$,*

$$\|m(K_{p_1} - K_{p_2})\|_{\mathcal{F}} \leqslant L_1 \|m\|_{\mathcal{F}} \|p_1 - p_2\|_{\mathrm{tv}},$$
$$\|\varpi(p_1) - \varpi(p_2)\|_{\mathcal{F}} \leqslant L_2 \|p_1 - p_2\|_{\mathrm{tv}}.$$

As we will see in Appendix B, if we ask for the first bound to be verified for all $m \in \mathcal{M}$, the set of all signed bounded measures on $(E, \mathcal{E})$ (on which $\|\cdot\|_{\mathcal{F}}$ is extended in the obvious way), then hypothesis 3.1 implies the second bound. Anyway, the latter inequality will be a consequence of a new condition (hypothesis 4.1) we will impose in § 4.

We still need a supplementary notation: if $(X_n, S_n)_{n\in\mathbb{N}}$ is a Markov chain on $E \times \mathcal{P}(E)$ constructed as in § 1, we denote for all $n \in \mathbb{N}$,

$$S_n^{\varpi} := \frac{1}{n+1} \sum_{0\leqslant i\leqslant n} \varpi(S_i).$$

Now our main result from this section can be stated as the following.

Proposition 3.3. *Under the hypotheses 3.1 and 3.2, for all $p \geqslant 1$, there exists a constant $C_p \geqslant 0$ such that, for all $n \in \mathbb{N}^*$,*

$$\sup_{f\in\mathcal{F}} \mathbb{E}[|S_n[f] - S_n^{\varpi}[f]|^p]^{1/p} \leqslant \frac{C_p}{\sqrt{n}}.$$

In the traditional case of just one mixing probability kernel, K, verifying the second condition of hypothesis 3.1, we have that, for all $n \in \mathbb{N}$, $S_n^{\varpi} = \varpi$, the unique invariant probability of K and the above result is a well-known bound for the Markovian law of large numbers.

As announced, the proof will be based on resolvent estimates. If a couple $(f, p) \in \mathcal{F} \times \mathcal{P}(E)$ is given, the associated resolvent $F_{f,p}$ is the function defined by

$$\forall x \in E, \quad F_{f,p}(x) = \sum_{n\in\mathbb{N}} (K_p^n[f](x) - \varpi(p)[f]).$$

To be convinced that this series is in fact absolutely convergent, we consider the term associated in the sum to a generic $n \in \mathbb{N}$, which we write $n = kn_0 + l$, with $k \in \mathbb{N}$ and $0 \leqslant l < n_0$. We then have

$$\begin{aligned} |K_p^n[f](x) - \varpi(p)[f]| &= |(\delta_x - \varpi(p))K_p^{kn_0+l}[f]| \\ &\leqslant \|(\delta_x - \varpi(p))K_p^{l+kn_0}\|_{\mathcal{F}} \\ &\leqslant \lambda^k \|(\delta_x - \varpi(p))K_p^l\|_{\mathcal{F}} \\ &\leqslant 2\lambda^k \end{aligned}$$

(one will have noticed that the second condition of hypothesis 3.1 can be rewritten as $\|mK_p^{n_0}\|_{\mathcal{F}} \leqslant \lambda \|m\|_{\mathcal{F}}$, for all $m \in \mathcal{M}_0(E)$ and all $p \in \mathcal{P}(E)$). This upper bound enables us to see that $F_{f,p} \in \mathcal{B}(E)$ and more precisely that $\|F_{f,p}\|_\infty \leqslant 2n_0/(1-\lambda)$. Note that we have not yet used the first condition of hypothesis 3.1, but it will be useful to deduce the main estimate needed to prove proposition 3.3, and that we now set forth.

Lemma 3.4. *Under hypothesis 3.1 and hypothesis 3.2, we have that, for all $f \in \mathcal{F}$ and all $p_1, p_2 \in \mathcal{P}(E)$,*

$$\|F_{f,p_1} - F_{f,p_2}\|_\infty \leqslant \left\{ 2\left(\frac{L_0^{n_0} - 1}{(L_0 - 1)(1-\lambda)}\right)^2 L_1 + \frac{L_0^{n_0} - 1}{(L_0 - 1)(1-\lambda)} L_2 \right\} \|p_1 - p_2\|_{\mathrm{tv}}$$

(as usual, if $L_0 = 1$, the fractions $(L_0^{n_0} - 1)/(L_0 - 1)$ should be understood as n_0).

Proof. Let us first admit that, for all $f \in \mathcal{F}$, for all $p_1, p_2 \in \mathcal{P}(E)$ and for all $x \in E$, we are assured of the equality

$$\begin{aligned} F_{f,p_1}(x) - F_{f,p_2}(x) + \varpi(p_1)[F_{f,p_2}] \\ = \sum_{n,q\in\mathbb{N}} (\delta_x - \varpi(p_1)) K_{p_1}^n (K_{p_1} - K_{p_2}) K_{p_2}^q [f - \varpi(p_2)[f]]. \end{aligned} \tag{3.1}$$

We then consider two generical natural integers, n, q, which we write $n = kn_0 + l$ and $q = rn_0 + s$, with $k, r \in \mathbb{N}$ and $0 \leqslant l, s < n_0$. We then have, for the corresponding term,

$$\begin{aligned} |(\delta_x - \varpi(p_1)) K_{p_1}^n (K_{p_1} - K_{p_2}) K_{p_2}^q [f - \varpi(p_2)[f]]| \\ &= |(\delta_x - \varpi(p_1)) K_{p_1}^n (K_{p_1} - K_{p_2}) K_{p_2}^q [f]| \\ &\leqslant \|(\delta_x - \varpi(p_1)) K_{p_1}^n (K_{p_1} - K_{p_2}) K_{p_2}^q \|_{\mathcal{F}} \\ &\leqslant \lambda^r \|(\delta_x - \varpi(p_1)) K_{p_1}^n (K_{p_1} - K_{p_2}) K_{p_2}^s \|_{\mathcal{F}} \\ &\leqslant \lambda^r L_0^s \|(\delta_x - \varpi(p_1)) K_{p_1}^n (K_{p_1} - K_{p_2}) \|_{\mathcal{F}} \\ &\leqslant \lambda^r L_0^s L_1 \|(\delta_x - \varpi(p_1)) K_{p_1}^n \|_{\mathcal{F}} \|p_1 - p_2\|_{\mathrm{tv}} \\ &\leqslant \lambda^{r+k} L_0^s L_1 \|(\delta_x - \varpi(p_1)) K_{p_1}^l \|_{\mathcal{F}} \|p_1 - p_2\|_{\mathrm{tv}} \\ &\leqslant \lambda^{r+k} L_0^{s+l} L_1 \|\delta_x - \varpi(p_1)\|_{\mathcal{F}} \|p_1 - p_2\|_{\mathrm{tv}} \\ &\leqslant 2\lambda^{r+k} L_0^{s+l} L_1 \|p_1 - p_2\|_{\mathrm{tv}}. \end{aligned}$$

This computation enables us on one hand to see that the right-hand side of (3.1) is absolutely convergent and on the other hand to deduce the upper bound presented in the above lemma by summation, via the estimate

$$\begin{aligned} |\varpi(p_1)[F_{f,p_2}]| &= \left| \varpi(p_1) \left[\sum_{n\in\mathbb{N}} (K_{p_2}^n[f] - \varpi(p_2)) \right] \right| \\ &= \left| \sum_{n\in\mathbb{N}} (\varpi(p_1) - \varpi(p_2)) [K_{p_2}^n[f]] \right| \\ &\leqslant \sum_{n\in\mathbb{N}} \|(\varpi(p_1) - \varpi(p_2)) K_{p_2}^n \|_{\mathcal{F}} \\ &\leqslant (1 + L_0 + \cdots + L_0^{n_0-1}) \sum_{k\in\mathbb{N}} \lambda^k \|\varpi(p_1) - \varpi(p_2)\|_{\mathcal{F}} \\ &\leqslant \frac{L_0^{n_0} - 1}{L_0 - 1} \frac{1}{1-\lambda} L_2 \|p_1 - p_2\|_{\mathrm{tv}}. \end{aligned}$$

So it remains to verify the validity of formula (3.1). But by the absolute convergence of the right-hand side, it is sufficient to show that

$$F_{f,p_1}(x) - F_{f,p_2}(x) + \varpi(p_1)[F_{f,p_2}] = \lim_{N\to\infty} A_N(f, p_1, p_2, x),$$

where

$$A_N(f,p_1,p_2,x) := \sum_{0\leqslant n,q\leqslant N} (\delta_x - \varpi(p_1))K_{p_1}^n(K_{p_1}-K_{p_2})K_{p_2}^q[f-\varpi(p_2)[f]]$$
$$= (\delta_x - \varpi(p_1))\bigg(\sum_{0\leqslant n,q\leqslant N} K_{p_1}^n(K_{p_1}-K_{p_2})K_{p_2}^q\bigg)[f].$$

But we remark that in the above sum there are a lot of cancellations and in fact

$$\begin{aligned}
&A_N(f,p_1,p_2,x)\\
&= (\delta_x - \varpi(p_1))\bigg(K_{p_1}^{N+1}\sum_{0\leqslant q\leqslant N} K_{p_2}^q - \sum_{0\leqslant n\leqslant N} K_{p_1}^n K_{p_2}^{N+1}\\
&\qquad + \sum_{0\leqslant n\leqslant N} K_{p_1}^{n+1}K_{p_2}^0 - K_{p_1}^0\sum_{0\leqslant q\leqslant N} K_{p_2}^{q+1}\bigg)[f]\\
&= (\delta_x - \varpi(p_1))\bigg(K_{p_1}^{N+1}\sum_{0\leqslant q\leqslant N} K_{p_2}^q - \sum_{0\leqslant n\leqslant N} K_{p_1}^n K_{p_2}^{N+1}\\
&\qquad + \sum_{-1\leqslant n\leqslant N} K_{p_1}^{n+1} - \sum_{-1\leqslant q\leqslant N} K_{p_2}^{q+1}\bigg)[f]\\
&= -(\delta_x - \varpi(p_1))\bigg(\sum_{0\leqslant n\leqslant N} K_{p_1}^n K_{p_2}^{N+1}[f] + \sum_{0\leqslant q\leqslant N+1} K_{p_2}^q[f-\varpi(p_2)[f]]\bigg)\\
&\quad + (\delta_x - \varpi(p_1))\bigg(\sum_{0\leqslant q\leqslant N} K_{p_1}^{N+1}K_{p_2}^q[f] + \sum_{0\leqslant n\leqslant N+1} K_{p_1}^n[f-\varpi(p_1)[f]]\bigg).
\end{aligned}$$

But we have

$$\begin{aligned}
&(\delta_x - \varpi(p_1))\bigg|\sum_{0\leqslant n\leqslant N} K_{p_1}^n K_{p_2}^{N+1}[f]\bigg|\\
&\qquad\leqslant (1\vee L_0^{n_0-1})\lambda^{\lfloor N/n_0\rfloor}\bigg\|(\delta_x - \varpi(p_1))\sum_{0\leqslant n\leqslant N} K_{p_1}^n\bigg\|_{\mathcal{F}}\\
&\qquad\leqslant (1\vee L_0^{n_0-1})\lambda^{\lfloor N/n_0\rfloor}\frac{L_0^{n_0}-1}{L_0-1}\frac{1}{1-\lambda}\|\delta_x - \varpi(p_1)\|_{\mathcal{F}}\\
&\qquad\leqslant 2(1\vee L_0^{n_0-1})\lambda^{\lfloor N/n_0\rfloor}\frac{L_0^{n_0}-1}{L_0-1}\frac{1}{1-\lambda},
\end{aligned}$$

an expression which tends to zero as N goes to infinity. In a similar way we get

$$\lim_{N\to\infty}\bigg|(\delta_x - \varpi(p_1))\sum_{0\leqslant q\leqslant N} K_{p_1}^{N+1}K_{p_2}^q[f]\bigg| = 0,$$

so, taking into account the definition of the resolvents, we obtain

$$\begin{aligned}
\lim_{N\to\infty} A_N(f,p_1,p_2,x) &= -(\delta_x - \varpi(p_1))[F_{f,p_2}] + (\delta_x - \varpi(p_1))[F_{f,p_1}]\\
&= F_{f,p_1}(x) - F_{f,p_2}(x) + \varpi(p_1)[F_{f,p_2}].
\end{aligned}$$

∎

The main interest in resolvent $F_{f,p}$ is due to the fact that it is a solution $F \in \mathcal{B}(E)$ of the Poisson equation

$$\left.\begin{aligned} K_p[F] - F &= -(f - \varpi(p)[f]),\\ \varpi(p)[F] &= 0, \end{aligned}\right\} \tag{3.2}$$

as it is checked at once (one could alternatively have used this property to verify (3.1), replacing the factor $K_{p_1} - K_{p_2}$ by $K_{p_1} - \mathrm{Id} - (K_{p_2} - \mathrm{Id})$).

Remark 3.5. We are not sure that, under our general setting, there is a unique solution to this equation. Hypothesis 3.1 implies there is at most one solution belonging to $\mathcal{F}_p$, the closure of $\mathrm{Vect}(\bigcup_{n\in\mathbb{N}} K_p^n(\mathcal{F}))$ with respect to the uniform norm. So, for instance, if we know that K_p maps $\mathcal{B}(E)$ into the previous vector space $\mathcal{F}_p$ (in particular, if $\mathcal{F} = \{f \in \mathcal{B}(E) : \|f\|_\infty \leqslant 1\}$), then we can conclude to the uniqueness of the solution F. In fact, in the cases where $\|\cdot\|_{\mathcal{F}}$ is equivalent to $\|\cdot\|_{\mathrm{tv}}$, one also gets the uniqueness property for the solution of (3.2) directly from hypothesis 3.1. Nevertheless, we note that, in the general situation, one easily deduces from the uniqueness of the invariant probability $\varpi(p)$ that two solutions of the Poisson equation are at least $\varpi(p)$-a.s. equal.

We have now at our disposal all the ingredients necessary for the proof of proposition 3.3. The function $f \in \mathcal{F}$ being fixed, we consider the process M defined by

$$\forall n \in \mathbb{N}, \quad M_n = \sum_{m=0}^{n-1} F_{f,S_m}(X_{m+1}) - K_{S_m}[F_{f,S_m}](X_m)$$

(by a traditional convention, $M_0 = 0$). Taking into account (1.1), it is clear that M is a martingale with respect to the filtration $(\mathcal{T}_n)_{n\in\mathbb{N}}$, where, for all $n \in \mathbb{N}$, $\mathcal{T}_n$ is the σ-algebra generated by $X_0, X_1, \dots, X_n$. Its non-predictable quadratic variation process is given by

$$\forall n \in \mathbb{N}, \quad \langle M\rangle_n := \sum_{m=0}^{n-1} (F_{f,S_m}(X_{m+1}) - K_{S_m}[F_{f,S_m}](X_m))^2,$$

so via the Burkholder–Davis–Gundy inequality for discrete time martingales (see, for instance, Shiryaev 1996) and the uniform bound on the resolvent, we get that, for all $p \geqslant 1$, there exists a constant $c(p) > 0$ such that

$$\begin{aligned} \forall n \in \mathbb{N}, \quad \mathbb{E}\biggl[\sup_{0\leqslant m\leqslant n-1} |M_m|^p\biggr]^{1/p} &\leqslant c(p)\mathbb{E}[\langle M\rangle_n^{p/2}]^{1/p} \\ &\leqslant 2c(p)\biggl(\sum_{0\leqslant m\leqslant n-1} \sup_{p\in\mathcal{P}(E)} \|F_{f,p}\|_\infty^2\biggr)^{1/2} \\ &\leqslant \frac{2n_0 c(p)}{1-\lambda}\sqrt{n}. \end{aligned}$$

Thus, denoting $C(p) = 2n_0 c(p)/(1-\lambda)$, we have that, for all $n \in \mathbb{N}$,

$$\mathbb{E}[|M_n|^p] \leqslant C(p) n^{p/2} \tag{3.3}$$

But we note that M is related to the process $(S_n[f] - S_n^{\varpi}[f])_{n\in\mathbb{N}}$, since we have, for all $n \in \mathbb{N}$,

$$
\begin{aligned}
(n+1)(S_n[f] - S_n^{\varpi}[f]) &= \sum_{0\leqslant m\leqslant n} f(X_m) - \varpi(S_m)[f] \\
&= \sum_{0\leqslant m\leqslant n} F_{f,S_m}(X_m) - K_{S_m}[F_{f,S_m}](X_m) \\
&= \sum_{0\leqslant m\leqslant n} F_{f,S_m}(X_{m+1}) - K_{S_m}[F_{f,S_m}](X_m) \\
&\quad + \sum_{0\leqslant m\leqslant n} F_{f,S_{m+1}}(X_{m+1}) - F_{f,S_m}(X_{m+1}) \\
&\quad + \sum_{0\leqslant m\leqslant n} F_{f,S_m}(X_m) - F_{f,S_{m+1}}(X_{m+1}) \\
&= M_{n+1} + \sum_{0\leqslant m\leqslant n} \{F_{f,S_{m+1}}(X_{m+1}) - F_{f,S_m}(X_{m+1})\} \\
&\quad + F_{f,S_0}(X_0) - F_{f,S_{n+1}}(X_{n+1}).
\end{aligned}
$$

Lemma 3.4 enables us to evaluate the intermediate sum by showing there is a constant $C \geqslant 0$ such that

$$
\begin{aligned}
&\left| \sum_{0\leqslant m\leqslant n} F_{f,S_{m+1}}(X_{m+1}) - F_{f,S_m}(X_{m+1}) \right| \\
&\qquad \leqslant \sum_{0\leqslant m\leqslant n} \|F_{f,S_{m+1}} - F_{f,S_m}\|_{\infty} \\
&\qquad \leqslant C \sum_{0\leqslant m\leqslant n} \|S_{m+1} - S_m\|_{\mathrm{tv}} \\
&\qquad \leqslant C \sum_{0\leqslant m\leqslant n} \frac{1}{m+2} + (m+1)\left(\frac{1}{m+1} - \frac{1}{m+2}\right) \\
&\qquad \leqslant C \sum_{0\leqslant m\leqslant n} \frac{2}{m+2} \\
&\qquad \leqslant 2C \int_1^{n+2} \frac{1}{t}\,\mathrm{d}t \\
&\qquad = 2C\ln(n+2).
\end{aligned}
$$

Thus, we obtain that there exists a constant $C' \geqslant 0$ such that, for all $n \in \mathbb{N}$,

$$
|(n+1)(S_n[f] - S_n^{\varpi}[f]) - M_{n+1}| \leqslant C'(1 + \ln(2+n))
$$

and proposition 3.1 follows immediately from the upper bound (3.3).

4. Almost-sure convergence

The title of this section is a little misleading, as we will not prove that under appropriate conditions, $\mathbb{P}$-a.s. S_n converges $\mathcal{F}$-weakly to $\varPi$, but rather that, for any $f \in \mathcal{F}$,

$\mathbb{P}$-a.s. $S_n[f]$ converges to $\Pi[f]$. Of course there is no difference from the former affirmation if $\mathcal{F}$ contains an at most denumerable subset determining the $\mathcal{F}$-weak convergence.

Our basic tool will be to first obtain convenient bounds on $\mathbb{L}^p$ convergence, for $p \geqslant 2$, by using the result of § 3. But we will need to add the following hypothesis to hypotheses 3.1 and 3.2.

Hypothesis 4.1. *For any $(f, m) \in \mathcal{F} \times \mathcal{P}(E)$ there exists a non-negative measure $\gamma_{f,m}$ on the measurable space $(\mathcal{F}, \mathcal{E}(\mathcal{F}))$ satisfying*

$$\Lambda := \sup\{\gamma_{f,m}(\mathbb{1}) : f \in \mathcal{F},\, m \in \mathcal{P}(E)\} < 1$$

and such that, for any $f \in \mathcal{F}$, we are assured of the following inequalities for all $p_1, p_2 \in \mathcal{P}(E)$,

$$|\varpi(p_1)[f] - \varpi(p_2)[f]| \leqslant \int |p_1[h] - p_2[h]|\gamma_{f,p_1}(\mathrm{d}h).$$

Here, as previously, for all $p \in \mathcal{P}(E)$, $\varpi(p)$ is the invariant probability of K_p, whose existence and uniqueness is ensured by hypothesis 3.1. We note that by considering suprema over $f \in \mathcal{F}$, hypothesis 4.1 implies a strict contraction property for the mapping $\varpi : \mathcal{P}(E) \to \mathcal{P}(E)$, from which one deduces, as in § 2 *a*, the existence and uniqueness of the fixed probability Π. We also get that the second condition of hypothesis 3.2 is satisfied with $L_2 = \Lambda < 1$, due to the general bound $\|\cdot\|_{\mathcal{F}} \leqslant \|\cdot\|_{\mathrm{tv}}$.

This new hypothesis 4.1 reminds us of that of § 2 and indeed it is more general, since we will verify in Appendix B that hypothesis 2.1 implies hypothesis 4.1 with $\Lambda = \lambda_1/(1 - \lambda_2)$ under a mild measurability condition.

Nevertheless, here we do not want to bother with logarithmic factors which flourish naturally in upper bounds as in theorem 2.2, so we will say that a sequence $(a_n)_{n\in\mathbb{N}}$ of positive reals is at most of rate $r \in \mathbb{R}$ if

$$\limsup_{n\to\infty} \frac{\ln(a_n)}{\ln(n)} \leqslant r$$

(in particular, a negative rate ensures the convergence of the sequence to zero). With this terminology, our main step in the proof of a.s. convergence can be as stated in the following.

Proposition 4.2. *Under hypotheses 3.1, 3.2 and 4.1, we have that, for any fixed $p \geqslant 1$, the sequence $(\sup_{f\in\mathcal{F}} \mathbb{E}[|S_n[f] - \Pi[f]|^p]^{1/p})_{n\in\mathbb{N}}$ is at most of rate $-r$ with $r := (\frac{1}{2}) \wedge (1 - \Lambda) > 0$.*

At least for $p = 2$, this estimate is compatible with theorem 2.2.

Our goal is now a simple consequence of this result.

Theorem 4.3. *Under the conditions of proposition 4.2, we have for any fixed $f \in \mathcal{F}$, the $\mathbb{P}$-a.s. convergence of $S_n[f]$ towards $\Pi[f]$ for large n.*

Proof. Proposition 4.2 shows in particular that, for any fixed $p \geqslant 1$ and $f \in \mathcal{F}$, the sequence $(\mathbb{E}[|S_n[f] - \Pi[f]|^p])_{n\in\mathbb{N}}$ is at most of rate $-pr$, so if $p > 1/r$ this sequence is summable. Thus, if $\epsilon > 0$ is given, we have for such a fixed $p > r$,

$$\sum_{n\in\mathbb{N}} \mathbb{P}[|S_n[f] - \Pi[f]| > \epsilon] \leqslant \frac{1}{\epsilon^p} \sum_{n\in\mathbb{N}} \mathbb{E}[|S_n[f] - \Pi[f]|^p] < +\infty,$$

which implies by the Borel–Cantelli lemma that $\mathbb{P}$-a.s. $\limsup_{n\to\infty} |S_n[f]-\Pi[f]| \leqslant \epsilon$. As the parameter $\epsilon > 0$ can be arbitrary small, we conclude to the validity of the above theorem. ■

So it remains to prove the bounds of proposition 4.2. Due to proposition 3.3, which shows that, for any given $p \geqslant 1$, $(\sup_{f\in\mathcal{F}} \mathbb{E}[|S_n[f] - S_n^{\varpi}[f]|^p]^{1/p})_{n\in\mathbb{N}}$ is at most of rate $-(\frac{1}{2})$, it is sufficient to see that $(\sup_{f\in\mathcal{F}} \mathbb{E}[|S_n^{\varpi}[f] - \Pi[f]|^p]^{1/p})_{n\in\mathbb{N}}$ is at most of rate $-r$. With the help of Hölder's inequalities, we only need to consider the case where $p \in \mathbb{N}^*$. This observation enables us to proceed by a recurrence argument; assume that, for some fixed $p \in \mathbb{N}^*$, $(\sup_{f\in\mathcal{F}} \mathbb{E}[|S_n^{\varpi}[f] - \Pi[f]|^p]^{1/p})_{n\in\mathbb{N}}$ is at most of rate $-r$ and let us check it for $p+1$. We first remark that $(\sup_{f\in\mathcal{F}} \mathbb{E}[|S_n^{\varpi}[f] - \Pi[f]|^{p'}])_{n\in\mathbb{N}}$ is then at most of rate $-rp'$, for any $0 \leqslant p' \leqslant p$. We denote, for any $f \in \mathcal{F}$ and $n \in \mathbb{N}$,

$$\begin{aligned} I_n^{(p+1)}[f] &:= \mathbb{E}[|(n+1)(S_n^{\varpi}[f] - \Pi[f])|^{p+1}], \\ I_n^{(p+1)} &:= \sup_{f\in\mathcal{F}} I_n^{(p+1)}[f]. \end{aligned}$$

We have to verify that $(I_n^{(p+1)})_{n\in\mathbb{N}}$ is at most of rate $(1-r)(p+1)$.

In the same spirit as in the proof of theorem 2.2, we consider the expansion

$$\begin{aligned} I_{n+1}^{(p+1)}[f] &= \mathbb{E}\biggl[\biggl|\sum_{0\leqslant i\leqslant n+1} \varpi(S_i)[f] - \Pi[f]\biggr|^{p+1}\biggr] \\ &\leqslant \sum_{0\leqslant k\leqslant p+1} \binom{p+1}{k} \mathbb{E}[|(n+1)(S_n^{\varpi}[f] - \Pi[f])|^{p+1-k}|\varpi(S_{n+1})[f] - \Pi[f]|^k]. \end{aligned} \tag{4.1}$$

By our iteration assumption, it appears that the sequence

$$\biggl(\sum_{2\leqslant k\leqslant p+1} \binom{p+1}{k} \mathbb{E}[|(n+1)(S_n^{\varpi}[f] - \Pi[f])|^{p+1-k}|\varpi(S_{n+1})[f] - \Pi[f]|^k]\biggr)_{n\in\mathbb{N}} \tag{4.2}$$

is at most of rate $(1-r)(p-1)$.

Thus, the more interesting term in inequality (4.1) is certainly the second one (after $I_n^{(p+1)}[f]$),

$$(p+1)\mathbb{E}[|(n+1)(S_n^{\varpi}[f] - \Pi[f])|^p|\varpi(S_{n+1})[f] - \Pi[f]|],$$

whose absolute value is bounded via hypothesis 4.1 by

$$
\begin{aligned}
(p+1)\mathbb{E}[(n+1)^p|S_n^{\varpi}[f]-\Pi[f]|^p|\varpi(S_{n+1})[f]-\varpi(\Pi)[f]|]&\\
\leqslant (p+1)\mathbb{E}\bigg[(n+1)^p|S_n^{\varpi}[f]-\Pi[f]|^p\int|S_{n+1}[h]-\Pi[h]|\gamma_{f,\Pi}(\mathrm{d}h)\bigg]&\\
\leqslant (p+1)\int\mathbb{E}[(n+1)^p|S_n^{\varpi}[f]-\Pi[f]|^p|S_{n+1}[h]-S_{n+1}^{\varpi}[h]|]\gamma_{f,\Pi}(\mathrm{d}h)&\\
+(p+1)\int\mathbb{E}[(n+1)^p|S_n^{\varpi}[f]-\Pi[f]|^p|S_{n+1}^{\varpi}[h]-S_n^{\varpi}[h]|]\gamma_{f,\Pi}(\mathrm{d}h)&\\
+(p+1)\int\mathbb{E}[(n+1)^p|S_n^{\varpi}[f]-\Pi[f]|^p|S_n^{\varpi}[h]-\Pi[h]|]\gamma_{f,\Pi}(\mathrm{d}h).&
\end{aligned}
$$

We consider separately each term on the right-hand side of the last inequality.

Resorting again to Hölder's inequality, the latter term is bounded by

$$\frac{p+1}{n+1}\int(I_n^{(p+1)}[f])^{p/(p+1)}(I_n^{(p+1)}[h])^{1/(p+1)}\gamma_{f,\Pi}(\mathrm{d}h)\leqslant\Lambda\frac{p+1}{n+1}I_n^{(p+1)}.$$

Due to the observation that

$$\|S_{n+1}^{\varpi}-S_n^{\varpi}\|_{\mathrm{tv}}\leqslant 2/(n+1),$$

the intermediate term cannot be larger than $2\Lambda(p+1)I_n^{(p)}/(n+1)$, whose corresponding sequence is at most of rate $(1-r)p-1$ by our recurrence assumption.

In order to treat the first term, we note that, for any $\epsilon>0$, we can find a constant $A:=A(p,\epsilon)$ such that

$$\forall x,y\geqslant 0,\quad x^p y\leqslant\epsilon x^{p+1}+Ay^{p+1}.$$

Thus, we obtain that

$$
\begin{aligned}
&\int\mathbb{E}[(n+1)^p|S_n^{\varpi}[f]-\Pi[f]|^p|S_{n+1}[h]-S_{n+1}^{\varpi}[h]|]\gamma_{f,\Pi}(\mathrm{d}h)\\
&\qquad\leqslant\epsilon\int\mathbb{E}[(n+1)^p|S_n^{\varpi}[f]-\Pi[f]|^{p+1}]\gamma_{f,\Pi}(\mathrm{d}h)\\
&\qquad\qquad+A\int\mathbb{E}[(n+1)^p|S_{n+1}[h]-S_{n+1}^{\varpi}[h]|^{p+1}]\gamma_{f,\Pi}(\mathrm{d}h)\\
&\qquad\leqslant\epsilon\Lambda I_n^{(p+1)}[f]/(n+1)+A\Lambda C_p(n+1)^{(p-1)/2},
\end{aligned}
$$

where we have once more used proposition 3.3.

Finally, we have shown that, for all $n\in\mathbb{N}$,

$$I_{n+1}^{(p+1)}\leqslant\bigg(1+\frac{(p+1)(1+\epsilon)\Lambda}{n+1}\bigg)I_n^{(p+1)}+R_n,$$

where $(R_n)_{n\in\mathbb{N}}$ is a sequence whose rate is at most

$$\beta:=(1-r)(p-1)\vee((1-r)p-1)\vee\tfrac{1}{2}(p-1).$$

From the bound $r\leqslant 1/2$, we deduce that $\beta+1\leqslant(1-r)(p+1)$, so lemma A 2 of Appendix A enables us to see that $(I_n^{(p+1)})_{n\in\mathbb{N}}$ is at most of rate

$$\{(1+\epsilon)\Lambda\vee(1-r)\}(p+1).$$

But, as is customary, ϵ can be arbitrary small. Thus, we end up with the expected result, since $1-r=\Lambda\vee(\frac{1}{2})$.

Only the initialization of the recurrence is lacking, for example, with $p=2$. Nevertheless, on re-examining the above computations, there is no difficulty in being convinced that $(I_n^{(2)})_{n\in\mathbb{N}}$ is at most of rate $2(1-r)$ (just note that, for $p=1$, we have $\beta=0$, since the sequence (4.2) is trivially at most of rate zero and clearly this is also true for $(2\Lambda(p+1)I_n^{(1)}/(n+1))_{n\in\mathbb{N}}$).

5. An example of extension

The example we present here is rather academic in nature, as its 'real world' aspect is maybe not very convincing! Nevertheless, some of its features give a good illustration of the role which can be played by the choice of convergence dictated by the class $\mathcal{F}$ and the advantage of hypothesis 4.1 over its counterpart, hypothesis 2.1.

First we need to remark that our previous proofs can be extended to the following situation, where a family of semi-norms is considered instead of a unique norm. We assume that we have at our disposal a family $(\mathcal{F}_j)_{j\in\mathcal{J}}$ of subsets of

$$\mathbf{B}:=\{f\in\mathcal{B}(E):\|f\|_\infty\leqslant 1\}.$$

For each j in the arbitrary index set $\mathcal{J}$, we denote by $\|\cdot\|_j$ the semi-norm associated to $\mathcal{F}_j$ on $\mathcal{M}(E)$:

$$\forall m\in\mathcal{M}(E),\quad \|m\|_j=\sup_{f\in\mathcal{F}_j}|m[f]|.$$

The hypothesis of completion made on $\|\cdot\|_{\mathcal{F}}$ in § 1 is now replaced by the weaker condition that if $(p_n)_{n\in\mathbb{N}}$ is a sequence of probabilities such that, for any $f\in\mathcal{F}:=\bigcup_{j\in\mathcal{J}}\mathcal{F}_j$, $p_n[f]$ is converging for large $n\in\mathbb{N}$, then there is a unique probability $p\in\mathcal{P}(E)$ verifying that the corresponding limits can be written as $p[f]$ (in particular, $\mathcal{F}$ separates $\mathcal{P}(E)$). This hypothesis will enable us to relax the uniform exponential convergence to invariant probabilities we have encountered in the above sections.

Remark 5.1. The resort to even more general systems of semi-norms could also be useful in situations where we are not assured of the uniqueness of the fixed point Π, as it is usual in traditional reinforced random walk context (see, for example, Benaïm 1997; Pemantle & Volkov 1999; Tarrès 2004).

Naturally, we replace hypotheses 3.1, 3.2 and 4.1, respectively, by the following.

Hypothesis 5.2. *For all $j\in\mathcal{J}$, there exist $L_0(j)\geqslant 0$, $n_0(j)\in\mathbb{N}^*$ and $0\leqslant\lambda(j)<1$ such that for all $p,m_1,m_2\in\mathcal{P}(E)$ we are assured of*

$$\|m_1K_p-m_2K_p\|_j\leqslant L_0(j)\|m_1-m_2\|_j,$$
$$\|m_1K_p^{n_0(j)}-m_2K_p^{n_0(j)}\|_j\leqslant\lambda(j)\|m_1-m_2\|_j.$$

There is then no difficulty in checking that, for any $p\in\mathcal{P}(E)$, there is a unique invariant probability $\varpi(p)$ for K_p.

Hypothesis 5.3. *For all $j\in\mathcal{J}$, there exist two constants $L_1(j),L_2(j)\geqslant 0$ such that, for all $m\in\mathcal{M}_0(E)$ and all $p_1,p_2\in\mathcal{P}(E)$,*

$$\|m(K_{p_1}-K_{p_2})\|_j\leqslant L_1(j)\|m\|_j\|p_1-p_2\|_{\mathrm{tv}},$$
$$\|\varpi(p_1)-\varpi(p_2)\|_j\leqslant L_2(j)\|p_1-p_2\|_{\mathrm{tv}}.$$

Hypothesis 5.4. *For any $j \in \mathcal{J}$ and $(f,m) \in \mathcal{F}_j \times \mathcal{P}(E)$ there exists a non-negative measure $\gamma_{f,m,j}$ on the measurable space $(\mathcal{F}, \mathcal{E}(\mathcal{F}))$ satisfying*

$$\Lambda(j) := \sup\{\gamma_{f,m,j}(\mathbb{1}) : f \in \mathcal{F}_j, m \in \mathcal{P}(E)\} < 1$$

and such that for any $(f,m) \in \mathcal{F}_j \times \mathcal{P}(E)$ we are assured of the following inequalities for all $p_1, p_2 \in \mathcal{P}(E)$:

$$|\varpi(p_1)[f] - \varpi(p_2)[f]| \leqslant \int |p_1[h] - p_2[h]| \gamma_{f,p_1,j}(\mathrm{d}h).$$

Then the previous proofs can be extended to prove the following theorem.

Theorem 5.5. *Under hypotheses 5.2–5.4, there exists a unique solution $\Pi \in \mathcal{P}(E)$ of the equation $\Pi = \Pi K_\Pi$ and we are assured, for any fixed $f \in \mathcal{F}$, of the $\mathbb{P}$-a.s. convergence of $S_n[f]$ towards $\Pi[f]$ for large n.*

Now, here is our example on $E = [0,1[^2$ endowed with $\mathcal{E}$, its traditional Borelian σ-algebra. We take $\mathcal{J} = \mathbb{Z}^2$ and, for any $j = (j_1, j_2) \in \mathbb{Z}^2$, we consider $\mathcal{F}_j := \{f_j, g_j\}$ with

$$\left.\begin{aligned} f_j(x) &= \cos(2\pi(j_1 x_1 + j_2 x_2)),\\ g_j(x) &= \sin(2\pi(j_1 x_1 + j_2 x_2)), \end{aligned}\right\} \quad \forall\, x = (x_1, x_2) \in E. \tag{5.1}$$

Indeed, the set $\mathcal{F}$ is here weak-convergence determining and so the first above-mentioned requirement is fulfilled.

Next we introduce our family of kernels $\mathbb{K}$. Let $r = (r_1, r_2) \in (\mathbb{R}^*)^2$ be such that r_1/r_2 is not a rational number, we consider $D = \mathrm{Vect}(r) \subset \mathbb{R}^2$ and Q the canonical projection from $\mathbb{R}^2$ to $\mathbb{R}^2/\mathbb{Z}^2$ identified with E. Note that due to the irrationality of r_1/r_2, the restriction of Q to D is an injection whose image is dense in E.

Let us also be given $\sigma : \mathcal{P}(E) \to [\sigma_1, \sigma_2] \subset]0, +\infty[$, a mapping which is Lipschitz for total variation norm; there exists a constant $M \geqslant 0$ such that

$$\forall\, p_1, p_2 \in \mathcal{P}(E), \quad |\sigma(p_1) - \sigma(p_2)| \leqslant M \|p_1 - p_2\|_{\mathrm{tv}}.$$

For any fixed $x \in E$ and $p \in \mathcal{P}(E)$, the probability $K_p(x, \cdot)$ can be described in the following way. Let $k_p(x, \cdot)$ be the measure on $x + D$ (x seen here as an element of $\mathbb{R}^2$) given by

$$k_p(x, \{x + tr : t \in A\}) = \int_A \exp\left(\frac{-t^2}{\{2\sigma(p)\}}\right) \frac{\mathrm{d}t}{\sqrt{2\pi\sigma(p)}}$$

for any Borelian subset A of $\mathbb{R}$. Then $K_p(x, \cdot)$ is the image by Q of $k_p(x, \cdot)$.

Let us check the above hypotheses. For any $j \in \mathcal{J}$, we can take, in hypothesis 5.2, $n_0(j) = 1$ and $\lambda(j) = \exp(-2\pi^2 \langle j, r\rangle^2 \sigma_1) < 1$, since we compute that

$$\forall\, p \in \mathcal{P}(E), \quad K_p[\exp(2\pi \mathrm{i} \langle j, \cdot\rangle)](\cdot) = \exp(-2\pi^2 \langle j, r\rangle^2 \sigma(p)) \exp(2\pi \mathrm{i}\langle j, \cdot\rangle).$$

For any $p \in \mathcal{P}(E)$, it appears that the unique invariant probability of K_p is the (restriction to E) of the two-dimensional Lebesgue measure, which consequently we can denote by Π.

To check hypothesis 5.3, we note that, for any $a, s_1, s_2 > 0$, we have

$$|\exp(-as_1) - \exp(-as_2)| \leqslant a|s_1 - s_2|,$$

so we can take, for any $j \in \mathcal{J}$, $L_1'(j) := 2\pi^2\langle j, r\rangle^2 M$ in hypothesis B 2 in Appendix B, corresponding to hypothesis 5.3.

The fact that the mapping $\varpi : \mathcal{P}(E) \to \mathcal{P}(E)$ is constant also implies that hypothesis 5.4 is trivially verified, thus finally we get that $\mathbb{P}$-a.s., S_n converges weakly (in the classical sense) to Π. Of course this result is not very interesting, since it is not the simplest way to simulate according to the Lebesgue measure on E! But from a theoretical point of view, this simple example puts forth the role played by the test-functions sets $(\mathcal{F}_j)_{j\in\mathcal{J}}$ (in comparison with the ball $\mathbf{B}$, note that $S_n[f]$ is not converging to $\Pi[f]$ for any $f \in \mathcal{B}(E)$) and the advantage of hypothesis 5.4 over hypothesis 2.1 (or rather its obvious modification in the spirit of § 5), which is not necessarily satisfied here, especially if M is large.

Furthermore, clearly in this situation there is not a unique solution to the Poisson equation (3.2), since for instance $F = 0$ or F the indicator function of $Q(D)$ are solutions associated to the homogeneous equation. This suggests that maybe one cannot expect more than uniqueness almost sure in remark 3.5, at least it is the case under the above generalized framework.

Appendix A. Usual difference inequalities

We present here discrete versions of classical differential inequalities which have been frequently used in this paper.

Lemma A 1. *Let $(y_n)_{n\in\mathbb{N}}$ be a sequence of non-negative numbers satisfying, for all $n \in \mathbb{N}$ large enough, the inequality*

$$y_{n+1} \leqslant \left(1 + \frac{\alpha}{n+1}\right) y_n + k(n+1)^\beta,$$

where $\alpha, \beta \in \mathbb{R}$ and $k > 0$ are given. There then exists a constant $c > 0$ such that, for all $n \in \mathbb{N}^* \setminus \{1\}$, we are assured of the upper bounds

$$\begin{aligned} \alpha - \beta < 1 &\Rightarrow y_n \leqslant cn^{\beta+1}, \\ \alpha - \beta = 1 &\Rightarrow y_n \leqslant cn^{\alpha}\ln(n), \\ \alpha - \beta > 1 &\Rightarrow y_n \leqslant cn^{\alpha}. \end{aligned}$$

Proof. By analogy with the continuous case, we just consider the sequence $(z_n)_{n\in\mathbb{N}}$ defined by

$$\forall\, n \in \mathbb{N}, \quad z_n = \frac{y_n}{a_n}$$

with $a_n = \prod_{0\leqslant i\leqslant n}(1 + \alpha/(1+i))$. Its interest is due to the fact that we have, for all $n \in \mathbb{N}$ with $n \geqslant n_0$ and $n_0 \in \mathbb{N}$ fixed large enough,

$$\begin{aligned} z_{n+1} &\leqslant z_n + k\frac{(n+1)^\beta}{a_{n+1}} \\ &\leqslant z_{n_0} + k\sum_{n_0\leqslant j\leqslant n} \frac{(j+1)^\beta}{a_{j+1}}. \end{aligned}$$

The announced estimates are now a straightforward consequence of the well-known fact that there exist two constants $0 < c_1 < c_2$ such that, for all $n \in \mathbb{N}^*$, we have

$$c_1(n+1)^\alpha \leqslant a_n \leqslant c_2(n+1)^\alpha.$$

■

The same arguments also show that if $(y_n)_{n\in\mathbb{N}}$ is a sequence of non-negative numbers satisfying, for all $n \in \mathbb{N}$ large enough, the inequality

$$y_{n+1} \geqslant \left(1 + \frac{\alpha}{n+1}\right) y_n + k(n+1)^\beta,$$

where $\alpha, \beta \in \mathbb{R}$ and $k > 0$ are given, then there exists a constant $c > 0$ such that, for all $n \in \mathbb{N}^* \setminus \{1\}$, we have

$$\begin{aligned}
\alpha - \beta < 1 &\Rightarrow y_n \geqslant cn^{\beta+1},\\
\alpha - \beta = 1 &\Rightarrow y_n \geqslant cn^{\alpha}\ln(n),\\
\alpha - \beta > 1 &\Rightarrow y_n \geqslant cn^{\alpha}.
\end{aligned}$$

This remark enables us to simplify the proof considered in Del Moral & Miclo (2002) relative to the example mentioned after theorem 2.2. In fact, via Hölder's inequalities, this example can also be used to see that the rates derived in proposition 4.2 are optimal in the presented setting, at least for $p \geqslant 2$. But let us show that they can be of the right order even for $p = 1$. In this case, the example corresponds to the situation where, for all $p \in \mathcal{P}(E)$, the kernel K_p is indeed the probability $\varpi(p)$ given by $\Lambda p + (1-\Lambda)\Pi$, where $0 < \Lambda < 1$ and $\Pi \in \mathcal{P}(E)$ are, respectively, a fixed parameter and a given probability, coinciding with the notions designated similarly in the rest of the paper. Let $f \in \mathcal{B}(E)$ be any function. Then one has, for any $n \in \mathbb{N}^*$,

$$\begin{aligned}
(n+1)\mathbb{E}[|S_n[f] - \Pi[f]|] &= \mathbb{E}[\mathbb{E}[|f(X_n) - \Pi[f] + n(S_{n-1}[f] - \Pi[f])||\mathcal{T}_n]]\\
&\geqslant \mathbb{E}[|\mathbb{E}[f(X_n) - \Pi[f] + n(S_{n-1}[f] - \Pi[f])|\mathcal{T}_n]|]\\
&= \mathbb{E}[|\mathbb{E}[K_{S_{n-1}}[f](X_{n-1}) - \Pi[f] + n(S_{n-1}[f] - \Pi[f])]|]\\
&= \left(1 + \frac{\Lambda}{n}\right) n\mathbb{E}[|S_{n-1}[f] - \Pi[f]|].
\end{aligned}$$

This lower bound implies immediately that the sequence $(\mathbb{E}[|S_n[f] - \Pi[f]|])_{n\in\mathbb{N}}$ is at least of order $\Lambda - 1$. Thus, in the case where $\frac{1}{2} \leqslant \Lambda < 1$, proposition 4.2 leads to the conclusion that

$$\lim_{n\to\infty} \frac{\ln(\mathbb{E}[|S_n[f] - \Pi[f]|])}{\ln(n)} = -(1-\Lambda).$$

In other respects, one cannot dispense with the term $\frac{1}{2}$ in proposition 4.2, as shown by classical i.i.d. examples (with no interaction at all, namely, for all $p \in \mathcal{P}(E)$, $K_p = \Pi$).

Another useful variant of lemma A 1 enables us to reinterpret its difference inequality as a transformer of rate.

Lemma A 2. *Let $(y_n)_{n\in\mathbb{N}}$ be a sequence of non-negative numbers satisfying, for all $n\in\mathbb{N}$ large enough, the inequality*

$$y_{n+1}\leqslant\left(1+\frac{\alpha}{n+1}\right)y_n+a_n,$$

where $(a_n)_{n\in\mathbb{N}}$ is a sequence of positive reals, the rate of which is at most β, with $\alpha,\beta\in\mathbb{R}$. $(y_n)_{n\in\mathbb{N}}$ is then at most of rate $\alpha\vee(\beta+1)$.

Proof. By definition, for any fixed $\epsilon>0$, we can find a constant $A>0$ such that, for all $n\in\mathbb{N}$ large enough, $a_n\leqslant A(n+1)^{\beta+1}$ (in fact we can take, for instance, $A=1$). Applying lemma A 1 and letting ϵ go to zero then gives the announced result. ∎

Appendix B. About the hypotheses

We will discuss here the different relations between our conditions.

As shown by the example considered in § 5, in some situations hypothesis 4.1 can be easier to check than hypothesis 2.1! Nevertheless, when we have no *a priori* information on $\varpi:\mathcal{P}(E)\to\mathcal{P}(E)$, hypothesis 2.1 gives at least one access to hypothesis 4.1, if we furthermore assume that, for any fixed $m\in\mathcal{P}(E)$ and $F\in\mathcal{B}(\mathcal{F})$, the mappings

$$\mathcal{F}\ni f\mapsto\mu_{f,m}[F]\in\mathbb{R},$$
$$\mathcal{F}\ni f\mapsto\nu_{f,m}[F]\in\mathbb{R}$$

are $\mathcal{E}(\mathcal{F})$-measurable.

To be convinced of the assertion that hypothesis 2.1 then implies hypothesis 4.1, we come back to the computations of § 2 *a*, but, instead of just deriving consequences of hypothesis 2.1 relative to contraction estimates for $d_{\mathcal{F}}$, we rather reconsider them in the spirit of this hypothesis. We begin by introducing for all $(f,m)\in\mathcal{F}\times\mathcal{P}(E)$ a sequence $(\mu_{f,m}^{(n)})_{n\in\mathbb{N}}$ of measures on $(\mathcal{F},\mathcal{E}(\mathcal{F}))$, defined through the iterations

$$\forall\ n\in\mathbb{N},\quad \mu_{f,m}^{(n+1)}(\cdot)=\int\mu_{f,m}(\mathrm{d}h)\mu_{h,m}^{(n)}(\cdot)$$

(taking into account the first above assumption of measurability), starting with $\mu_{f,m}^{(0)}=\delta_f$ (in particular, $\mu_{f,m}^{(1)}=\mu_{f,m}$).

It is relatively immediate to verify that, for all $n\in\mathbb{N}$ and all $(f,m)\in\mathcal{F}\times\mathcal{P}(E)$,

$$\forall p_1,p_2\in\mathcal{P}(E),\quad |p_1K_m^n[f]-p_2K_m^n[f]|\leqslant\int|p_1(h)-p_2(h)|\mu_{f,m}^{(n)}(\mathrm{d}h)$$

and

$$\mu_{f,m}^{(n)}(\mathbb{1})\leqslant\lambda_1^n.$$

Now it is sufficient to consider, for all $f\in\mathcal{F}$ and all $m\in\mathcal{P}(E)$,

$$\gamma_{f,m}(\cdot):=\sum_{n\in\mathbb{N}}\int\mu_{f,m}^{(n)}(\mathrm{d}h)\nu_{h,m}(\cdot),$$

since the relation

$$\forall f \in \mathcal{F},\ \forall p_1, p_2 \in \mathcal{P}(E), \quad \varpi(p_1)[f] - \varpi(p_2)[f] = \sum_{n\in\mathbb{N}} \varpi(p_2)(K_{p_1} - K_{p_2})K_{p_1}^n[f]$$

enables us to conclude that hypothesis 4.1 is satisfied by the family

$$(\gamma_{f,m})_{(f,m)\in\mathcal{F}\times\mathcal{P}(E)} \quad \text{and} \quad \varLambda = \lambda_1/(1-\lambda_2).$$

Let us now check the assertion made on condition hypothesis 3.2 after its statement.

As can be seen by reconsidering the second computation made in § 2 *a* to prove (2.2), hypothesis 3.2 is a consequence of hypothesis 3.1 and the following hypothesis.

Hypothesis B 1. *There exists a constant* $L_1' \geqslant 0$ *such that, for all* $m \in \mathcal{M}_0(E) \cup \mathcal{P}(E)$ *and all* $p_1, p_2 \in \mathcal{P}(E)$,

$$\|m(K_{p_1} - K_{p_2})\|_{\mathcal{F}} \leqslant L_1' \|m\|_{\mathcal{F}} \|p_1 - p_2\|_{\mathrm{tv}}.$$

(We can then take $L_1 = L_1'$ and $L_2 = (L_0^{n_0} - 1)(L_0 - 1)^{-1}(1-\lambda)^{-1}L_1'$ in hypothesis 3.2.)

So the simplest way to deduce hypothesis 3.2 is to verify the following hypothesis (and in practice, it seems difficult to check the first condition of hypothesis 3.2 without also deriving B 2!).

Hypothesis B 2. *There exists a constant* $L_1'' \geqslant 0$ *such that, for all* $m \in \mathcal{M}(E)$ *and all* $p_1, p_2 \in \mathcal{P}(E)$,

$$\|m(K_{p_1} - K_{p_2})\|_{\mathcal{F}} \leqslant L_1'' \|m\|_{\mathcal{F}} \|p_1 - p_2\|_{\mathrm{tv}}.$$

In some cases, hypotheses B 1 and B 2 are in fact equivalent, for instance if $\|\cdot\|_{\mathcal{F}} = \|\cdot\|_{\mathrm{tv}}$. This assertion comes from the observation that in this situation, for all $m \in \mathcal{M}$, the quantity $\|m\|_{\mathrm{tv}}$ is equal to

$$\min\{\|m_0\|_{\mathrm{tv}} + \|p\|_{\mathrm{tv}} : m = m_0 + p,\ \text{with } m_0 \in \mathcal{M}_0,\ p \in \mathrm{Vect}(\mathcal{P}(E))\},$$

as can be shown from the Hahn–Jordan decomposition of measures.

But we are not sure this equivalence of hypotheses B 1 and B 2 is still true for general test-functions collections $\mathcal{F}$.

References

Benaïm, M. 1997 Vertex-reinforced random walks and a conjecture of Pemantle. *Ann. Prob.* **25**, 361–392.

Del Moral, P. & Miclo, L. 2002 Markov chains with self-interactions. Technical report from the Laboratoire de Statistique et Probabilités, Université Paul Sabatier (Toulouse 3), France.

Pemantle, R. & Volkov, S. 1999 Vertex-reinforced random walk on $\mathbb{Z}$ has finite range. *Ann. Prob.* **27**, 1368–1388.

Shiryaev, A. N. 1996 *Probability*, 2nd edn. Springer. (Transl. R. P. Boas (1980) Russian 1st edn.)

Tarrès, P. 2004 Vortex-reinforced random walk on $\mathbb{Z}$ eventually gets stuck at a set of five points. *Ann. Prob.* (In the press.)