stochastic processes and their applications

Stochastic Processes and their Applications 86 (2000) 193–216

www.elsevier.com/locate/spa

# A Moran particle system approximation of Feynman–Kac formulae

P. Del Moral *, L. Miclo

*UMR C5583, CNRS, Université Toulouse III France*

Received 29 October 1998; received in revised form 30 September 1999

**Abstract**

We present a weighted sampling Moran particle system model for the numerical solving of a class of Feynman–Kac formulae which arise in different fields. Our major motivation was from nonlinear filtering, but our approach is context free. We will show that under certain regularity conditions the resulting interacting particle scheme converges to the considered nonlinear equations. In the setting of nonlinear filtering, the $\mathbb{L}^1$-convergence exponent resulting from our proof also improves recent results on other particle interpretations of these equations. © 2000 Elsevier Science B.V. All rights reserved.

*Keywords:* Weighted sampling Moran processes; Measure valued dynamical systems defined by Feynman–Kac formulae; Quantitative weak propagation of chaos; Nonlinear filtering

## 1. Introduction

### 1.1. Motivations

Initially, our motivation comes from nonlinear filtering. Heuristically, its purpose is to guess the position at some time $t$ of a Markovian signal process, given some undirect and noisy observations made up to this time $t$. The usual way to solve this problem is to compute the relative conditional expectations. If the observation process is a uniformly elliptic diffusion and if the signal merely acts on its drift, it is well known, via several uses of Girsanov formula, that the conditional distributions under interest can be easily written in terms of renormalized Feynman–Kac formulae.

But for the practitioner this theoretical solution is difficult to manipulate, since the Feynman–Kac formulae involve integration over a set of trajectories which is a very large space.

Our purpose here is to solve numerically these formulae by approximating them via interacting particle systems. It seems that the algorithm we propose is the first one which is genuinely continuous time, i.e. without any resort to discretization.

* Correspondence address: Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 118, route de Narbonne, 31062 Toulouse cedex, France.

*E-mail addresses:* delmoral@cict.fr (P. Del Moral), miclo@cict.fr (L. Miclo)

0304-4149/00/$ - see front matter 
PII: S0304-4149(99)00094-0

The basic mechanism behind its evolution is a Moran interaction, which is traditional in measure valued process theory as a particle approximation of the Fleming–Viot process (cf. Dawson (1993)). But the latter is stochastic, whereas we are looking for a deterministic measure valued dynamical system in the limit. We will manage to get it, by modifying the renormalization and by using a nonsymmetrical weight. In fact, our process should rather be seen as a Nanbu particle system approximation of a particular spatially homogeneous generalized Boltzmann equation (cf. Graham and Méléard, 1997).

But our approach will be different from the proofs arising in the literature of this field, and will take into account the specificity of our model. So the convergence analysis will rather be based on martingales and semi-group technics. May be our semi-group method can be extended to more general particle models whose interaction is expressed through the jumps, as those presented in Graham and Méléard (1997).

The paper has the following structure. In Section 2 we introduce some basic objects and terminology and discuss the hypotheses needed for further developments. In Section 3 we consider the infinitesimal pregenerator associated to the Moran particle scheme. The study of the weak propagation of chaos is performed in Section 4.

In the last section, we present several examples of mutation pregenerators that can be handled in our framework including bounded generators, Riemannian and Euclidean diffusions.

But we will not discuss here about the applications to nonlinear filtering problems. We just point out that in the more traditional situations, the signals are also such diffusions, that is why we have treated these examples with some details, showing how our abstract hypotheses retranscribe in this set-up.

### 1.2. Description of our Moran particle model and statement of some results

On a Polish space $E$, we assume that we are given two time inhomogeneous and measurable families $U=(U_t)_{t\geqslant 0}$ and $L=(L_t)_{t\geqslant 0}$, respectively, of measurable, non-negative and bounded functions and of pregenerators. Let us denote by $\boldsymbol{M}_1(E)$ the set of all probability measures on $E$, and we fix some element $\eta_0\in\boldsymbol{M}_1(E)$. Let $(X_t)_{t\geqslant 0}$ be a Markovian process whose initial law is $\eta_0$ and whose family of pregenerators is $L$. The object of interest in this article is the measure valued dynamical system defined by the renormalized Feynman–Kac formulae:

$$\forall t\geqslant 0,\quad \eta_t(f)=\frac{\mathbb{E}[f(X_t)\exp(\int_0^t U_s(X_s)\,\mathrm{d}s)]}{\mathbb{E}[\exp(\int_0^t U_s(X_s)\,\mathrm{d}s)]},$$

where $f$ is a measurable bounded function on $E$. It appears that $(\eta_t)_{t\geqslant 0}$ is a particular solution of the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta_t(f)=\eta_t(L_tf)+\eta_t(fU_t)-\eta_t(f)\eta_t(U_t). \tag{1}$$

We get formally the nature of our interacting particle schemes by noting that, for regular functions $f$, (1) can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\eta_t(f)=\eta_t(\mathscr{L}_{t,\eta_t}(f)),$$

where $\mathscr{L}_{t,\eta}$, for $t \geqslant 0$ and $\eta \in \boldsymbol{M}_1(E)$ fixed, is a pregenerator on $E$, defined on a suitable domain by

$$\mathscr{L}_{t,\eta}(f)(x) = L_t(f)(x) + \int (f(z) - f(x)) U_t(z) \eta(\mathrm{d}z). \tag{2}$$

Starting from this formula, we consider an interacting $N$-particles system $(\xi_t^{(N)})_{t \geqslant 0} = ((\xi_t^{(N,1)}, \dots, \xi_t^{(N,N)}))_{t \geqslant 0}$, which is a time-inhomogeneous Markov process on the product space $E^N$, $N \geqslant 1$, whose pregenerator $\mathscr{L}_t^{(N)}$ at time $t \geqslant 0$ acts on functions $\phi$ belonging to a good domain by

$$\mathscr{L}_t^{(N)}(\phi)(x_1, \dots, x_N) = \sum_{i=1}^{N} \mathscr{L}_{t,m^{(N)}(x)}^{(i)}(\phi)(x_1, \dots, x_N) \tag{3}$$

with

$$\forall x = (x_1, \dots, x_N) \in E^N, \quad m^{(N)}(x) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \in \boldsymbol{M}_1(E), \tag{4}$$

where $\delta_a$ stands for the Dirac measure at $a \in E$ and where the notation $\mathscr{L}_{t,\eta}^{(i)}$ have been used instead of $\mathscr{L}_{t,\eta}$ when it acts on the $i$th variable of $\phi(x_1, \dots, x_N)$. Assume that the initial particle system $\xi_0^{(N)} = (\xi_0^{(N,1)}, \dots, \xi_0^{(N,N)})$ consists of $N$-independent random variables with common law $\eta_0$. The main purpose of this work is to show that the empirical distributions of the $N$-particle system

$$\eta_t^{(N)} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\xi_t^{(N,i)}}$$

weakly converges as $N \to \infty$ to the desired solution of (1) and to give an upper bound on the speed.

**Theorem 1.1.** *Under certain regularity conditions, for any $T \geqslant 0$ and for any nice test function $f$,*

$$\sup_{0 \leqslant t \leqslant T} \mathbb{E}(|\eta_t^{(N)}(f) - \eta_t(f)|) \leqslant \frac{C_T}{\sqrt{N}} |||f||| \tag{5}$$

*for some constant $C_T$ which only depends on the time parameter $T$ (in particular through certain quantities associated with the underlying Markov semigroup, this dependence will be later more explicit), and where the norm $|||\cdot|||$ will be explained in Section 2.*

In nonlinear filtering settings the pregenerator of the particle scheme will use the observation record and the quenched version of (5) holds although the constant $C_T$ will also depend on the observations up to time $T$, because $U_t$ and $L_t$ depend on the observation at time $t \geqslant 0$. Note that for this kind of application, the time-inhomogeneous assumption is crucial due to the fact that we cannot ask for much regularity in time (typically $U_t$ and $L_t$ would not be differentiable with respect to time $t \geqslant 0$, and this forbids us from using the traditional trick of considering the homogeneous process $(t, X_t)_{t \geqslant 0}$, since the time–space function $U$ will not belong to the domain of its generator). In this framework the same scheme can also be used to approximate the optimal filter.

## 2. Hypotheses and preliminary results

To describe precisely our model, let us introduce some notations. Let $E$ be a Polish space, endowed with its Borelian structure. We denote by $\boldsymbol{M}(E)$ the space of all nonnegative and finite Borel measures on $E$. We will also designate by $\mathscr{C}_{\mathrm{b}}(E)$ (resp. $\mathscr{B}_{\mathrm{b}}(E)$) the Banach space of all bounded continuous functions (resp. of all bounded measurable functions), both sets being endowed with the supremum norm $\| \cdot \|$.

We now need an $E$-valued time-inhomogeneous Markov process $X = \{X_t;\ t \geqslant 0\}$. There are several ways of giving such an object, but perhaps the more convenient is via martingales problems (cf. Ethier and Kurtz, 1986, for a general reference).

Let $\mathscr{A}$ be a dense sub-algebra of $\mathscr{C}_{\mathrm{b}}(E)$ which is supposed to contain $\mathbb{1}$, the function taking everywhere the value 1.

Let $(L_t)_{t\geqslant 0}$ be a measurable family of pregenerators from the domain $\mathscr{A}$ to $\mathscr{C}_{\mathrm{b}}(E)$: for each $t \geqslant 0$, $L_t : \mathscr{A} \to \mathscr{C}_{\mathrm{b}}(E)$ is a linear operator satisfying the maximum principle (for the definition of this property, see for instance Proposition 2.2, p. 13 of Liggett (1985)) and such that $L_t(\mathbb{1}) \equiv 0$, and for each $f \in \mathscr{A}$ fixed,

$$\mathbb{R}_+ \times E \ni (t, x) \mapsto L_t(f)(x)$$

is $\mathscr{B}(\mathbb{R}_+) \otimes \mathscr{E}$-measurable, where $\mathscr{B}(\mathbb{R}_+)$ (respectively $\mathscr{E}$) is the $\sigma$-algebra of the Borelian sets on $\mathbb{R}_+$ (resp. on $E$). To get rid of integrability problems, we will also impose that for all $f \in \mathscr{A}$ and all $T > 0$,

$$\int_0^T \| L_t(f) \| \, \mathrm{d}t < +\infty.$$

For $t \geqslant 0$, let $D([t, +\infty[, E)$ be the set of all càdlàg paths from $[t, +\infty[$ to $E$, and we denote by $(X_s)_{s\geqslant t}$ the process of canonical coordinates on $D([t, +\infty[, E)$, which generate on this space the $\sigma$-algebra $\mathscr{D}_{t,+\infty} = \sigma(X_s \colon s \geqslant t)$.

Our first hypothesis is:

(H1) *For all $(t, x) \in \mathbb{R}_+ \times E$, there exists a unique probability $\mathbb{P}_{t,x}$ on $(D([t, +\infty[\, , E), \mathscr{D}_{t,+\infty})$ such that*

- *$X_t \circ \mathbb{P}_{t,x} = \delta_x$, the Dirac mass in $x$, and*
- *for all $f \in \mathscr{A}$, the process*

$$\left( f(X_s) - f(X_t) - \int_t^s L_u(f)(X_u) \, \mathrm{d}u \right)_{s \geqslant t}$$

*is a $(\mathscr{D}_{t,s})_{s\geqslant t}$-martingale under $\mathbb{P}_{t,x}$.*

Let us first precise that the previous martingales problem can be extended to a time–space version:

Let $\boldsymbol{A}$ be the set of absolutely continuous functions $g : \mathbb{R}_+ \to \mathbb{R}$ admitting a bounded derivative, i.e. there exists a bounded measurable function $g' : \mathbb{R}_+ \to \mathbb{R}$, such that for all $t \geqslant 0$,

$$g(t) = g(0) + \int_0^t g'(s) \, \mathrm{d}s.$$

On $\boldsymbol{A}\otimes\mathscr{A}$, we define the operator $L$ given on functions of the form $f=g\otimes h$, with $g\in\boldsymbol{A}$ and $h\in\mathscr{A}$, by

$$\forall t\geqslant 0,\quad \forall x\in E,\qquad L(f)(t,x)=g'(t)h(x)+g(t)L_t(h)(x).$$

Then we have:

**Lemma 2.1.** *Let $(t,x)\in\mathbb{R}_+\times E$ be fixed. Under $\mathbb{P}_{t,x}$, for each $f\in\boldsymbol{A}\otimes\mathscr{A}$, the process $(M_s(f))_{s\geqslant t}$ defined by*

$$\forall s\geqslant t,\qquad M_s(f)=f(s,X_s)-f(t,X_t)-\int_t^s L(f)(u,X_u)\,\mathrm{d}u$$

*is a square integrable martingale and its increasing process has the form*

$$\forall s\geqslant t,\qquad \langle M(f)\rangle_s=\int_t^s \Gamma(f,f)(X_u,u)\,\mathrm{d}u,$$

*where $\Gamma$ is the "carré du champ" bilinear operator associated to the pregenerator $L$ and defined by*

$$\forall\phi,\varphi\in\boldsymbol{A}\otimes\mathscr{A},\qquad \Gamma(\phi,\varphi)=L(\phi\varphi)-\phi L(\varphi)-\varphi L(\phi). \tag{6}$$

We can consider, for $s\geqslant 0$, the "carré du champ" bilinear operator $\Gamma_s$ associated to the pregenerator $L_s$, which is naturally defined by

$$\forall f,g\in\mathscr{A},\qquad \Gamma_s(f,g)=L_s(fg)-fL_s(g)-gL_s(f)$$

and we easily check that for all $f\in\boldsymbol{A}\otimes\mathscr{A}$,

$$\forall(s,x)\in\mathbb{R}_+\times E,\qquad \Gamma(f,f)(s,x)=\Gamma_s(f(s,.),f(s,.))(x)$$

(so no derivability of $f$ on the time variable is required to define $\Gamma(f,f)$, and this fact will often be used below).

But $\boldsymbol{A}\otimes\mathscr{A}$ is quite too small for our purpose, so let us extend it in the following way: for $T>0$ fixed, we denote by $A(T,\mathscr{A})$ the set of functions $f:[0,T]\times E\to\mathbb{R}$ such that for all $0\leqslant t\leqslant T$, $f(t,\cdot)\in\mathscr{A}$ and for which there exists a sequence $(f_n)_{n\geqslant 0}$ of elements of $\boldsymbol{A}_T\otimes\mathscr{A}$ satisfying $\sup_{n\geqslant 0,\,s\in[0,T],\,x\in E}|L(f_n)(s,x)|<+\infty$, $f_n\rightharpoonup f$ and $\Gamma(f_n-f,f_n-f)\rightharpoonup 0$, where $\rightharpoonup$ stands for the bounded pointwise convergence on $[0,T]\times E$, and where $\boldsymbol{A}_T$ is the set of restrictions to $[0,T]$ of functions belonging to $\boldsymbol{A}$.

We will need some regularity conditions on the family of probabilities $(\mathbb{P}_{t,x})_{(t,x)\in\mathbb{R}_+\times E}$, and these are expressed in the following hypothesis:

(H2) *For all $T>0$ and $\varphi\in\mathscr{A}$ fixed, the application*

$$F_{T,\varphi}\;:\;[0,T]\times E\ni(t,x)\mapsto\mathbb{E}_{t,x}[\varphi(X_T)] \tag{7}$$

*belongs to $A(T,\mathscr{A})$.*

This property has a lot of interesting consequences:

- First, we get that the application $F_{T,\varphi}$ is measurable with respect to $\mathscr{B}([0,T])\otimes\mathscr{E}$, and it is not difficult to deduce from this fact, by using the right continuity of the trajectories, that the function

$$\triangle\times E\ni(t,s,x)\mapsto\mathbb{E}_{t,x}[\varphi(X_s)]$$

is measurable with respect to the natural $\sigma$-algebra $\mathscr{B}(\triangle)\otimes\mathscr{E}$, where $\triangle=\{(t,s)\in\mathbb{R}_+^2:\ 0\leqslant t\leqslant s\}$.

If $\eta_0\in M_1(E)$, $\mathbb{P}_{\eta_0}$ will denote the probability on $(D([0,+\infty[,E),\mathscr{D}_{0,+\infty})$ defined by

$$\forall A\in\mathscr{D}_{0,+\infty},\qquad \mathbb{P}_{\eta_0}(A)=\int_E \mathbb{P}_{0,x}(A)\eta_0(\mathrm{d}x).$$

It is the unique solution to the martingales problem associated to $(L_t)_{t\geqslant 0}$ whose initial law is $\eta_0$.

- The strongly continuous and positive inhomogeneous semi-group $(P_{s,t})_{0\leqslant s\leqslant t}$ on $\mathscr{B}_{\mathrm{b}}(E)$ associated with the transition probabilities of $X\stackrel{\text{def.}}{=}(X_t)_{t\geqslant 0}$ is defined by

$$\forall 0\leqslant s\leqslant t,\quad \forall\varphi\in\mathscr{B}_{\mathrm{b}}(E),\quad \forall x\in E,\qquad P_{s,t}(\varphi)(x)=\mathbb{E}_{s,x}[\varphi(X_t)],$$

where $\mathbb{E}_{s,x}$ is the expectation relative to $\mathbb{P}_{s,x}$.

The first assumption in the definition of $A(T,\mathscr{A})$ shows in fact that for all $0\leqslant s\leqslant t$, $\mathscr{A}$ is stable under $P_{s,t}$.

In particular, since $P_{s,t}$ is a contraction (on $\mathscr{B}_{\mathrm{b}}(E)$) and $\mathscr{A}$ is dense in $\mathscr{C}_{\mathrm{b}}(E)$, it follows that $\mathscr{C}_{\mathrm{b}}(E)$ is stable by the operators $P_{s,t}$, $0\leqslant s\leqslant t$, so the Markov process $(X_t)_{t\geqslant 0}$ is Fellerian.

- But the consequence of (H2) which really matters for us is the following one: let us remark that if we note for $T>0$ and $\varphi\in\mathscr{A}$,

$$\forall 0\leqslant t\leqslant T,\qquad N_t(T,\varphi)=P_{t,T}(\varphi)(X_t),$$

then the Markov property of $X$ also implies that $(N_t(T,\varphi))_{0\leqslant t\leqslant T}$ is a martingale.

The interest of (H2) is that it enables us to get the following informations about this martingale (which would be immediate, under stronger regularity assumptions):

**Lemma 2.2.** *The martingale $(N_t(T,\varphi))_{0\leqslant t\leqslant T}$ is a.s. càdlàg, and its increasing process is*

$$\forall 0\leqslant t\leqslant T,\qquad \langle N(T,\varphi)\rangle_t=\int_0^t \Gamma_s(P_{s,T}(\varphi),P_{s,T}(\varphi))(X_s)\,\mathrm{d}s$$

**Proof.** Let $(f_n)_{n\geqslant 0}$ be a sequence of functions of $A_T\otimes\mathscr{A}$ corresponding to $F_{T,\varphi}$, in the sense of the above definition. From the general inequality

$$\Gamma(f_n-f_m,f_n-f_m)\leqslant 2(\Gamma(f_n-F_{T,\varphi},f_n-F_{T,\varphi})+\Gamma(f_m-F_{T,\varphi},f_m-F_{T,\varphi}))$$

valid for all $n,m\geqslant 0$, we obtain (via the application of a dominated convergence theorem) that

$$\lim_{n,m\to\infty}\mathbb{E}_{\eta_0}[(M_T(f_n)-M_T(f_m))^2]=\lim_{n,m\to\infty}\mathbb{E}_{\eta_0}[\langle M(f_n-f_m)\rangle_T]=0$$

and it is quite standard to deduce from this Cauchy convergence that there exists a martingale $(M_t)_{0\leqslant t\leqslant T}$ such that

$$\lim_{n\to\infty}\mathbb{E}_{\eta_0}\left[\sup_{0\leqslant t\leqslant T}(M_t(f_n)-M_t)^2\right]=0.$$

But we are also assurred, for $0 \leqslant t \leqslant T$, of the pointwise bounded convergence of $f_n(t,X_t) - f_n(0,X_0)$ toward $N_t(T,\varphi) - N_0(T,\varphi)$, so we have the pointwise convergence (may be only for a subsequence) of $\int_0^t L(f_n)(s,X_s)\,\mathrm{d}s$ toward $M_t - N_t(T,\varphi) - N_0(T,\varphi)$.

This shows that the latter process (for $0 \leqslant t \leqslant T$) is previsible, as a limit of previsible processes, that it has bounded variations (from our assumption on the uniform boundedness of the $L(f_n)$, $n \geqslant 0$), and as we already know that it is a martingale, we conclude that up to an evanescent set, it is null.

Then again an application of convergence theorems enables us to see that the process

$$\left([N_t(T,\varphi) - N_0(T,\varphi)]^2 - \int_0^t \Gamma_s(P_{s,T}(\varphi), P_{s,T}(\varphi))(X_s)\,\mathrm{d}s\right)_{0 \leqslant t \leqslant T}$$

is a martingale, from where the second affirmation of the lemma follows.

Note furthermore that hypothesis (H2) insures that $(N_t(T,\varphi))_{0 \leqslant t \leqslant T}$ is a.s. càdlàg, since it is equal to $(M_t)_{0 \leqslant t \leqslant T}$, which is càdlàg as a locally uniform (in time) limit of càdlàg martingales. □

This property leads us to extend the definition of $L$: for $T > 0$ fixed, let $D(T,L)$ be the vector space generated by $A_T \otimes \mathscr{A}$ and by $\{F_{T,\varphi}, F_{T,\varphi}^2 \colon \varphi \in \mathscr{A}\}$.

We agree to set, for all $\varphi \in \mathscr{A}$,

$\forall 0 \leqslant t \leqslant T$, $\forall x \in E$,

$$L(F_{T,\varphi})(t,x) = 0,$$

$$L(F_{T,\varphi}^2)(t,x) = \Gamma_t(P_{t,T}(\varphi), P_{t,T}(\varphi))(x).$$

Finally, we will also need in the sequel the following more quantitative assumption:

(H3) *There exists a convex subset* $\mathscr{D} \subset \mathscr{A}$ *with the following properties*:

- $\mathbb{1} \in \mathscr{D}$, $\mathscr{A} = \bigcup_{n \geqslant 1} n\mathscr{D}$ (*so we have* $0 \in \mathscr{D}$), *and* $\| \varphi \| \leqslant 1$ *for any* $\varphi \in \mathscr{D}$.
- *For any time* $T > 0$, *there exist three constants* $C_T^{(1)}$, $C_T^{(2)}$, $C_T^{(3)} < \infty$, increasing in $T$, such that for any $0 \leqslant t \leqslant T$ and any $\varphi \in \mathscr{D}$,

$$U_t\varphi \in C_T^{(1)}\,\mathscr{D},$$

$$P_{t,T}(\varphi) \in C_T^{(2)}\,\mathscr{D},$$

$$\| \Gamma_t(P_{t,T}(\varphi), P_{t,T}(\varphi)) \| \leqslant C_T^{(3)}$$

(*in particular, we will have* $U_t \in C_T^{(1)}\mathscr{D}$, *for* $0 \leqslant t \leqslant T$, *and so for all* $t \geqslant 0$, $U_t \in \mathscr{B}_\mathrm{b}(E)$, *and note that the three requirements above are still satisfied if* $T$ *is replaced by* $s$ *in the left-hand sides, with* $0 \leqslant t \leqslant s \leqslant T$)

Denote by $|||\cdot|||$ the jauge of $\mathscr{D}$ in $\mathscr{A}$:

$$\forall \varphi \in \mathscr{A}, \qquad |||\varphi||| = \inf\{l > 0 \colon \varphi/l \in \mathscr{D}\}$$

The convexity and the first property of $\mathscr{D}$ enable us to see that $|||\cdot|||$ is in fact a norm on $\mathscr{A}$ (the one alluded to in the introduction), which satisfies

$$\forall \varphi \in \mathscr{A}, \qquad |||\varphi||| \geqslant \| \varphi \| .$$

**Remark 2.3.** This norm can be different from the one usually put on $\mathscr{A}$, relatively to the interval $[0,T]$, and which is defined by

$$\forall \varphi \in \mathscr{A}, \qquad \| \varphi \|_{\mathscr{A},T} = \| \varphi \| + \sup_{0 \leqslant t \leqslant T} \| L_t \varphi \|$$

under the extra assumption that for all $\varphi \in \mathscr{A}$, $\| \varphi \|_{\mathscr{A},T} < +\infty$.

An example of more explicit domain $\mathscr{D}$ and related jauge $|||\cdot|||$ will be given later, see the first part of Section 5.

As all our hypotheses on the underlying Markov process $X$ are now put forward, we can describe the problem we are interested in. Let a probability $\eta_0 \in \boldsymbol{M}_1(E)$ be given, and consider a measurable application $U : \mathbb{R}_+ \times E \ni (t,x) \mapsto U(t,x) \in \mathbb{R}_+$, locally bounded, in the sense that for all given $T \geqslant 0$, the restriction of $U$ on $[0,T] \times E$ is bounded. The measure valued dynamical system under study is defined by the following Feynman–Kac formulae, for $t \geqslant 0$ and $f \in \mathscr{B}_b(E)$,

$$\gamma_t(f) \stackrel{\text{def}}{=} \mathbb{E}_{\eta_0}[f(X_t) e^{\int_0^t U_s(X_s)\,\mathrm{d}s}], \tag{8}$$

$$\eta_t(f) \stackrel{\text{def}}{=} \gamma_t(f)/\gamma_t(\mathbb{1}),$$

where for a time $s \geqslant 0$ fixed, $U_s$ denote the bounded measurable application $E \ni x \mapsto U(s,x) \in \mathbb{R}_+$.

These Feynman–Kac formulae are commonly used as a probabilistic representation for solutions of certain parabolic differential equations (see for instance Krylov (1964) and Sznitman (1997) and references therein) and it also plays a major role in nonlinear filtering theory.

In view of the functional representation (8) the temptation is to apply classical Monte-Carlo simulations based on a sequence of independent copies of the process $X$. Unfortunately, it is well known that the resulting particle scheme is not efficient mainly because the deviation of the particles may be too large and the growth of the exponential weights with respect to the time parameter is difficult to control.

This is not astonishing: roughly speaking the law of $X_t$ and the desired distribution $\eta_t$ may differ considerably and there may be too few particles in the space regions with height $\eta_t$-mass probability. In contrast to the latter the Moran particle approximating model involve the use of a system of particles which evolve in correlation with each other and give birth to offsprings depending on the fitness function $U$. This guarantees an occupation of the probability space regions proportional to their probability mass thus providing a stochastic grid which is related to the fitness function $U$.

Let us present the two evolution equations that will be used in the foregoing development.

**Proposition 2.4.** *The measure valued process* $\{\eta_t;\ t \geqslant 0\}$ *satisfies the following two integral equations, for all* $t \geqslant 0$:

$\forall f \in \mathscr{A}$,

$$\eta_t(f) = \eta_0(f) + \int_0^t \eta_s(L_s(f))\,\mathrm{d}s + \int_0^t [\eta_s(fU_s) - \eta_s(f)\eta_s(U_s)]\,\mathrm{d}s, \tag{9}$$

$\forall f \in \mathscr{B}_{\mathrm{b}}(E)$,

$$\eta_t(f) = \eta_0(P_{0,t}(f)) + \int_0^t [\eta_s(U_s P_{s,t}(f)) - \eta_s(P_{s,t}(f))\eta_s(U_s)]\,\mathrm{d}s. \tag{10}$$

*Furthermore, there exists a unique solution of* (10) *for arbitrary initial conditions* $\eta_0 \in \boldsymbol{M}_1(E)$.

**Proof.** Since in the Radon–Nykodim sense, we have a.s. for $t \geqslant 0$,

$$\eta_t(U_t) = \frac{\mathrm{d}}{\mathrm{d}t} \log \mathbb{E}_{\eta_0}\left[\exp \int_0^t U_s(X_s)\,\mathrm{d}s\right],$$

we obtain

$$\forall f \in \mathscr{B}_{\mathrm{b}}(E), \quad \eta_t(f) = \mathbb{E}_{\eta_0}[f(X_t)\mathrm{e}^{\int_0^t \tilde{U}_s(X_s)\,\mathrm{d}s}], \tag{11}$$

where for all $s \geqslant 0$ and all $x \in E$, we have defined $\tilde{U}_s(x) = U_s(x) - \eta_s(U_s)$.

Now, writing that for $f \in \mathscr{A}$, there exists a martingale $M^{(f)} = (M_t^{(f)})_{t \geqslant 0}$ such that for all $t \geqslant 0$,

$$f(X_t) = f(X_0) + \int_0^t L_s(f)(X_s)\,\mathrm{d}s + M_t^{(f)}$$

and that

$$\mathrm{e}^{\int_0^t \tilde{U}_s(X_s)\,\mathrm{d}s} = 1 + \int_0^t \tilde{U}_s(X_s)\,\mathrm{e}^{\int_0^s \tilde{U}_u(X_u)\,\mathrm{d}u}\,\mathrm{d}s, \tag{12}$$

we get by standard stochastic calculus that $\{\eta_t; t \geqslant 0\}$ is a solution of (9). To prove (10), we use one more time (12), because it yields that the right-hand side of (11) equals

$$\eta_0(P_{0,t}(f)) + \int_0^t \mathbb{E}_{\eta_0}[f(X_t)\tilde{U}_s(X_s)\,\mathrm{e}^{\int_0^s \tilde{U}_u(X_u)\,\mathrm{d}u}]\,\mathrm{d}s. \tag{13}$$

Using the Markov property of $X$, the last member of (13) is equal to

$$\int_0^t \mathbb{E}_{\eta_0}[P_{s,t}(f)(X_s)\tilde{U}_s(X_s)\,\mathrm{e}^{\int_0^s \tilde{U}_u(X_u)\,\mathrm{d}u}]\,\mathrm{d}s.$$

Again from (11), one concludes that

$$\begin{aligned}\eta_t(f) &= \eta_0(P_{0,t}(f)) + \int_0^t \eta_s(P_{s,t}(f)\tilde{U}_s)\,\mathrm{d}s \\ &= \eta_0(P_{0,t}(f)) + \int_0^t \eta_s((U_s - \eta_s(U_s))P_{s,t}(f))\,\mathrm{d}s\end{aligned}$$

and the proof of (10) is completed.

Let us check the uniqueness of the solution of (10). Let $\mathbb{R}_+ \ni t \mapsto \eta_t \in M_1(E)$ and $\mathbb{R}_+ \ni t \mapsto \bar{\eta}_t \in M_1(E)$ be measurable solutions of (10) with the same initial condition. We set

$$\forall t \geqslant 0,\ \forall f \in \mathscr{B}_{\mathrm{b}}(E), \quad I_t(f) = |\eta_t(f) - \bar{\eta}_t(f)|.$$

A direct computation yields that for any $t \geqslant 0$ and any $f \in \mathscr{B}_{\mathrm{b}}(E)$,

$$I_t(f) \leqslant \int_0^t [I_s(U_s P_{s,t}(f)) + \| U \|_t^\star I_s(P_{s,t}(f)) + \| f \| I_s(U_s)]\, \mathrm{d}s, \tag{14}$$

where $\| U \|_t^\star = \sup_{0 \leqslant s \leqslant t} \| U_s \|$. This implies that

$$\forall t \geqslant 0,\ \forall f \in \mathscr{B}_{\mathrm{b}}(E), \quad I_t(f) \leqslant t(3 \| U \|_t^\star) \| f \| . \tag{15}$$

Substituting (15) into the right-hand side of (14) we get

$$\forall t \geqslant 0,\ \forall f \in \mathscr{B}_{\mathrm{b}}(E), \quad I_t(f) \leqslant \frac{t^2}{2}(3 \| U \|_t^\star)^2 \| f \| .$$

Repeating this procedure $n$-times we arrive at

$$I_t(f) \leqslant \frac{t^n}{n!}(3 \| U \|_t^\star)^n \| f \| \underset{n \to \infty}{\longrightarrow} 0.$$

This ends the proof of the proposition. □

The evolution equation (9) will be used to define a pregenerator of the Moran interacting particle scheme. The second differential equation (10) gives a more tractable and general description of the desired valued measure process (8), at least once the inhomogeneous semi-group $(P_{s,t})_{0 \leqslant s \leqslant t}$ is known, fact which will always be assumed here. We will use this equation to study the convergence of such genetic-type interacting particle scheme.

## 3. The interacting particle system model

The genetic-type interacting particle system under study will be a Markov process $(\xi_t^{(N)})_{t \geqslant 0} = ((\xi_t^{(N,1)}, \ldots, \xi_t^{(N,N)}))_{t \geqslant 0}$ with state space $E^N$, where $N \geqslant 1$ is the size of the system.

Heuristically, the motion of the particles will be decomposed into the two following rules. Between the jumps due to interaction between particles, each particle evolves independently from the others and randomly according to a $L$-motion in $E$ (that is according to the time-inhomogeneous semigroup of $X$).

At some random times we introduce a competitive interaction beween the particles. More precisely, during this stage a chosen particle $\xi_t^{(N,i)}$ will be replaced by a new particle $\xi_t^{(N,j)}$, $1 \leqslant j \leqslant N$, with a probability proportional to its adaptation $U_t(\xi_t^{(N,j)})$, $1 \leqslant j \leqslant N$.

So at time $t \geqslant 0$ a pregenerator of the interacting particle scheme associated to (1) is the sum of two pregenerators

$$\forall t \geqslant 0, \qquad \mathscr{L}_t^{(N)} = \tilde{\mathscr{L}}_t^{(N)} + \hat{\mathscr{L}}_t^{(N)}.$$

The first pregenerator $\tilde{\mathscr{L}}_t^{(N)}$ is called the mutation pregenerator. It denotes the generator coming from $N$-independent $L$-processes and it is given on $\mathscr{A}^{\otimes N}$ by

$$\forall \phi \in \mathscr{A}^{\otimes N}, \quad \tilde{\mathscr{L}}_t^{(N)}(\phi)(x_1, \ldots, x_N) = \sum_{i=1}^N L_t^{(i)}(\phi)(x_1, \ldots, x_N),$$

where $L_t^{(i)}$ denotes the action of $L_t$ on the $i$th variable $x_i$, i.e. $L_t^{(i)} = \mathrm{Id} \otimes \cdots \otimes \underbrace{L_t}_{i\text{th}} \otimes \cdots \otimes \mathrm{Id}$, where Id is the identity operator.

The second one, $\hat{\mathscr{L}}_t^{(N)}$, is called the selection pregenerator. It is the jump-type generator defined by

$$\forall \phi \in \mathscr{A}^{\otimes N}, \quad \hat{\mathscr{L}}_t^{(N)}(\phi)(x_1,\ldots,x_N) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} \Delta_i^{x_j}(\phi)(x_1,\ldots,x_N)U_t(x_j),$$

where for $1 \leqslant i \leqslant N$ and $y \in E$,

$$\Delta_i^y(\phi)(x_1,\ldots,x_N) \stackrel{\text{def.}}{=} \phi(x_1,\ldots,\underbrace{y}_{i\text{th}},\ldots,x_N) - \phi(x_1,\ldots,x_N)$$

(this is meaningful for all functions $\phi \in \mathscr{C}_{\mathrm{b}}(E^N)$, and in fact $\hat{\mathscr{L}}_t^{(N)}$ is a bounded generator on $\mathscr{C}_{\mathrm{b}}(E^N)$).

For any fixed $N \geqslant 1$, the infinitesimal pregenerator $\mathscr{L}_t^{(N)}$ with domain the algebra with unity $\mathscr{A}^{\otimes N}$, can be regarded as the pregenerator of a Moran-type interacting particle scheme with competitive selection interactions (see for instance Dawson, 1993 and references therein). But in contrast to the classical Moran process the total jump rate is "proportional" to the number of particles, and above the roles of $i$ and $j$ are not symmetrical.

For any given probability $m_0$ on $E^N$, it is quite standard to construct explicitly a Markov process $(\xi_t^{(N)})_{t\geqslant 0} = ((\xi_t^{(N,1)},\ldots,\xi_t^{(N,N)}))_{t\geqslant 0}$ whose initial law is $m_0$ and verifying the martingale problem corresponding to the family of pregenerators $(\mathscr{L}_t^{(N)})_{t\geqslant 0}$. More precisely, using the fact that all $t \geqslant 0$, $\mathscr{L}_t^{(N)}$ is just a bounded perturbation of $\tilde{\mathscr{L}}_t^{(N)}$ by $\hat{\mathscr{L}}_t^{(N)}$, we can apply general results about this kind of martingales problems, see for instance the Proposition 10.2 p. 256 of Ethier and Kurtz (1986).

From now on, $\mathbb{E}$ will designate the expectation relative to the process $\xi^{(N)}$ defined in this section, starting with initial law $\eta_0^{\otimes N}$ at time 0.

## 4. Quantitative weak propagation of chaos results

In fact, the laws of the particle systems $(\xi_t^{(N)})_{0\leqslant t\leqslant T}$, where $N \in \mathbb{N}^*$ and $T > 0$ are fixed, satisfy more extended martingales problems, because the proofs presented by Ethier and Kurtz (1986) enable us to transpose the whole pregenerator $(D(L), L)$ considered in Section 2, owing to hypothesis (H2).

Let us denote by $A(T, N, \mathscr{A})$ the vector sub-space of $C_{\mathrm{b}}([0,T] \times E^N)$ generated by the functions $f(t,x) = \prod_{1\leqslant i\leqslant N} f_i(t,x_i)$, where $f_1,\ldots,f_N \in A(T,\mathscr{A})$.

If such a function $f$ is given, we define for all $(t,x) \in [0,T] \times E$,

$$\tilde{\mathscr{L}}^{(N)}(f)(t,x) = \sum_{1\leqslant i\leqslant N} f_1(t,x_1)\cdots f_{i-1}(t,x_{i-1})L(f_i)(t,x_i)f_{i+1}(t,x_{i+1})\cdots f_N(t,x_N)$$

and then we extend linearly this operator $\tilde{\mathscr{L}}^{(N)}$ on $A(T,N,\mathscr{A})$. We also consider the pregenerator $\hat{\mathscr{L}}^{(N)}$ acting on $A(T,N,\mathscr{A})$ in the following way:

$$\forall f \in A(T,N,\mathscr{A}), \forall (t,x) \in [0,T] \times E,$$

$$\hat{\mathscr{L}}_t^{(N)}(f)(t,x) = \hat{\mathscr{L}}_t^{(N)}(f(t,\cdot))(x)$$

and next we introduce

$$\mathscr{L}^{(N)} = \tilde{\mathscr{L}}^{(N)} + \hat{\mathscr{L}}^{(N)}.$$

This pregenerator on $A(T,N,\mathscr{A})$ coincides naturally with $\partial_t + \mathscr{L}_t^{(N)}$ on $\boldsymbol{A}_T \otimes \mathscr{A}^{\otimes N}$. Then we are assured of

**Lemma 4.1.** *Under the laws of $\xi^{(N)}$ constructed in the previous section, for all $T > 0$ and all $f \in A(T,N,\mathscr{A})$, the process $(M_t^{(N)}(f))_{0 \leqslant t \leqslant T}$ defined by*

$$\forall 0 \leqslant t \leqslant T, \quad M_t^{(N)}(f) = f(t,\xi_t^{(N)}) - f(0,\xi_0^{(N)}) - \int_0^t \mathscr{L}^{(N)}(f)(s,\xi_s^{(N)})\,\mathrm{d}s$$

*is a bounded martingale.*

We will apply this result to some special functions, for which $x \in E^N$ is seen only through its empirical measure $m^{(N)}(x)$ defined in (4).

Let $T > 0$ and $\varphi \in \mathscr{A}$ be fixed, we first consider the function $f_1 \in A(T,N,\mathscr{A})$ defined by

$$\forall 0 \leqslant t \leqslant T, \forall x \in E^N, \quad f_1(t,x) = m^{(N)}(x)[P_{t,T}(\varphi)] = \frac{1}{N} \sum_{1 \leqslant i \leqslant N} P_{t,T}(\varphi)(x_i).$$

One of its main interest is that it satisfies

$$\forall t \geqslant 0, \ \forall x \in E^N,$$

$$\tilde{\mathscr{L}}^{(N)}(f_1)(t,x) = 0,$$

$$\hat{\mathscr{L}}^{(N)}(f_1)(t,x) = m^{(N)}(x)[U_t P_{t,T}(\varphi)] - m^{(N)}(x)[U_t] m^{(N)}(x)[P_{t,T}(\varphi)].$$

Thus the process defined for $0 \leqslant t \leqslant T$ by

$$M_t^{(N)}(f_1) = \eta_t^{(N)}(P_{t,T}(\varphi)) - \eta_0^{(N)}(P_{0,T}(\varphi))$$
$$- \int_0^t (\eta_s^{(N)}(U_s P_{s,T}(\varphi)) - \eta_s^{(N)}(U_s)\eta_s^{(N)}(P_{s,T}(\varphi)))\,\mathrm{d}s$$

(let us recall that $\eta_t^{(N)} = m^{(N)}(\xi_t^{(N)})$) is a càdlàg martingale. To find its increasing process, let us consider the function $f_2 = f_1^2$. So we have

$$\forall 0 \leqslant t \leqslant T, \forall x \in E^N, \quad f_2(t,x) = \frac{1}{N^2} \sum_{1 \leqslant i,j \leqslant N} P_{t,T}(\varphi)(x_i) P_{t,T}(\varphi)(x_j),$$

where it clearly appears that this function belongs to $A(T,N,\mathscr{A})$.

We calculate that

$$\forall 0 \leqslant t \leqslant T, \forall x \in E^N,$$

$$\mathscr{L}^{(N)}(f_2)(t,x) = \frac{1}{N} m^{(N)}(x)[\Gamma_t(P_{t,T}(\varphi), P_{t,T}(\varphi))]$$
$$+ 2 f_1(t,x) \mathscr{L}^{(N)}(f_1)(t,x)$$
$$+ \frac{1}{N} m^{(N)}(x)[(P_{t,T}(\varphi) - m^{(N)}(x)[P_{t,T}(\varphi)])^2 (U_t + m^{(N)}(x)[U_t])].$$

We easily deduce from this fact that

$$(M_t^{(N)}(f_1))^2 - \frac{1}{N}\int_0^t (\eta_s^{(N)}[\Gamma_s(P_{s,T}(\varphi), P_{s,T}(\varphi))]$$
$$+ \eta_s^{(N)}[(P_{s,T}(\varphi) - \eta_s^{(N)}[P_{s,T}(\varphi)])^2 (U_s + \eta_s^{(N)}[U_s])])\, \mathrm{d}s$$

is a martingale (once again, this would be quite immediate, if one has at his disposal enough regularity of $[0,T] \times E^N \ni (t,x) \mapsto m^{(N)}(x)[P_{t,T}(\varphi)]$ in order to apply directly to this application and its square the time–space martingale problems for the $N$-particles system).

We are now in a position to prove the following weak propagation of chaos result with rate.

**Theorem 4.2.** *Under assumptions* (H1), (H2) *and* (H3) *we have that*

$$\sup_{0 \leqslant t \leqslant T} \sup_{\varphi \in \mathscr{D}} \mathbb{E}(|\eta_t^{(N)}(\varphi) - \eta_t(\varphi)|) \leqslant \frac{C_T}{\sqrt{N}} \tag{16}$$

*for some constant $C_T < \infty$ that only depends on the time parameter $T > 0$.*

**Proof.** The previous discussion easily implies that for any $0 \leqslant t \leqslant T$ and any $\varphi \in \mathscr{D}$, the process $B^{(N,t)}(\varphi) = (B_s^{(N,t)}(\varphi))_{0 \leqslant s \leqslant t}$ defined for $0 \leqslant s \leqslant t$ by

$$B_s^{(N,t)}(\varphi) = \eta_s^{(N)}(P_{s,t}(\varphi)) - \eta_0^{(N)}(P_{0,t}(\varphi)) - \int_0^s \eta_u^{(N)}(U_u P_{u,t}(\varphi))$$
$$- \eta_u^{(N)}(U_u) \eta_u^{(N)}(P_{u,t}(\varphi))\, \mathrm{d}u$$

is a bounded martingale and its increasing process has the form

$$\forall 0 \leqslant s \leqslant t,\ \langle B^{(N,t)}(\varphi) \rangle_s = \frac{1}{N} \int_0^s \tilde{F}_u(\eta_u^{(N)}, P_{u,t}(\varphi)) + \hat{F}_u(\eta_u^{(N)}, P_{u,t}(\varphi))\, \mathrm{d}u,$$

where for all $u \geqslant 0$, all $\eta \in \boldsymbol{M}_1(E)$ and all $\varphi \in \mathscr{A}$,

$$\tilde{F}_u(\eta, \varphi) = \eta(\Gamma_u(\varphi, \varphi)),$$
$$\hat{F}_u(\eta, \varphi) = \eta[(\varphi - \eta(\varphi))^2 (U_u + \eta(U))].$$

Under our assumptions we note that

$$\sup_{\varphi \in \mathscr{D}} \sup_{0 \leqslant u \leqslant t \leqslant T} \sup_{\eta \in \boldsymbol{M}_1(E)} \tilde{F}_u(\eta, P_{u,t}(\varphi)) + \hat{F}_u(\eta, P_{u,t}(\varphi)) \leqslant C_T \tag{17}$$

for some constant $C_T < \infty$ which could be explicited in terms of $C_T^{(1)}$, $C_T^{(2)}$ and $C_T^{(3)}$.

Now, we set

$$\forall 0 \leqslant t \leqslant T,\ \forall \varphi \in \mathscr{D}, \qquad I_t^{(N)}(\varphi) \stackrel{\text{def.}}{=} |\eta_t^{(N)}(\varphi) - \eta_t(\varphi)|.$$

Using the evolution equation (10), we prove that for any $0\leqslant t\leqslant T$ and any $\varphi\in\mathscr{D}$, the error bounds $I_t^{(N)}(\varphi)$ are bounded by

$$I_0^{(N)}(P_{0,t}(\varphi)) + \int_0^t (I_s^{(N)}(U_s P_{s,t}(\varphi)) + \parallel U \parallel_t^\star \; I_s^{(N)}(P_{s,t}(\varphi))$$

$$+ I_s^{(N)}(U_s))\,\mathrm{d}s + |B_t^{(N,t)}(\varphi)|,$$

where $\parallel U \parallel_t^\star = \sup_{0\leqslant s\leqslant t} \parallel U_t \parallel$. Note that (17) implies that for a new constant $C_T$,

$$\sup_{\varphi\in\mathscr{D}} \mathbb{E}\left(\sup_{0\leqslant t\leqslant T} |B_t^{(N,t)}(\varphi)|\right) \leqslant \frac{C_T}{\sqrt{N}}.$$

From our choice of the initial particle scheme let us also observe that for all $\varphi\in\mathscr{B}_\mathrm{b}(E)$,

$$\begin{aligned}\mathbb{E}(I_0^{(N)}(\varphi)) &\leqslant \sqrt{\mathbb{E}[(\eta_0^{(N)}[\varphi]-\eta_0(\varphi))^2]}\\ &\leqslant \frac{1}{\sqrt{N}}\sqrt{\eta_0[(\varphi-\eta_0(\varphi))^2]},\end{aligned}$$

so we have (recall that $\parallel P_{0,t}(\varphi) \parallel \leqslant \parallel \varphi \parallel \leqslant 1$ for $\varphi\in\mathscr{D}$)

$$\sup_{0\leqslant t\leqslant T} \sup_{\varphi\in\mathscr{D}} \mathbb{E}(I_0^{(N)}(P_{0,t}(\varphi))) \leqslant \frac{1}{\sqrt{N}}.$$

Then, under (H3) there exists an other constant $C_T>0$ such that for all $0\leqslant t\leqslant T$, if we define the error bound

$$\bar{I}_t^{(N)} \stackrel{\text{def.}}{=} \sup_{\varphi\in\mathscr{D}} \mathbb{E}[I_t^{(N)}(\varphi)],$$

then it is itself less than

$$C_T\left(\frac{1}{\sqrt{N}} + \int_0^t \bar{I}_s^{(N)}\,\mathrm{d}s\right)$$

and therefore (16) is now a clear consequence of Gronwall's lemma.

At least, this would be true if the application $[0,T]\ni t\mapsto \bar{I}_t^{(N)}$ is measurable (the fact that can be assured by assuming a pointwise separability of $\mathscr{D}$, which is not a very strong requirement, for instance $\{f\in\mathscr{B}_\mathrm{b}(E):\ \parallel f \parallel \leqslant 1\}$ is pointwise separable as a consequence of separability of the $\sigma$-algebra $\mathscr{E}$).

Otherwise, in order to verify the bound

$$\forall 0\leqslant t\leqslant T,\quad I_t^{(N)} \leqslant \frac{C_T\exp(C_T t)}{\sqrt{N}},$$

one just needs to be a little more careful but there is no real difficulty, by using a recursive method similar to that of the end of the proof of Proposition 2.4. □

**Remark 4.3.** The simplicity of the above proof makes it easy to generalize to similar situations, for instance, one can consider the setting of Graham and Méléard (1997) and Méléard (1996). Nevertheless this approach does not enable us to obtain strong propagation of chaos, in the sense that we get for instance the strong convergence of $(\xi_t^{(N,1)})_{t\geqslant 0}$ toward a suitably coupled non-linear process, whose family of laws is

$(\eta_t)_{t\geqslant 0}$ (a general reference for strong propagation of chaos is Sznitman (1991), but Sznitman is rather interested in interactions going through drifts, and he did not look to the case, as here, where the interactions between the particles are expressed through jumps). By an other method, based on interacting graphs, Graham and Méléard got this strong propagation in their situation of more general jump rates (taking place in $\mathbb{R}^d$, $d\geqslant 1$).

We can also use the particle density profiles $\{\eta_t^{(N)};\ t\geqslant 0\}$ to approximate the Feynman–Kac formula

$$\forall f\in\mathscr{C}_b(E),\quad \gamma_t(f)=\mathbb{E}_{\eta_0}(f(X_t)\,\mathrm{e}^{\int_0^t U_s(X_s)\,\mathrm{d}s}).$$

More precisely, if we put

$$\gamma_t^{(N)}(f)\stackrel{\text{def.}}{=}\mathrm{e}^{\int_0^t \eta_s^{(N)}(U_s)\,\mathrm{d}s}\eta_t^{(N)}(f).$$

**Corollary 4.4.** *Under the assumptions of Theorem* 4.2 *there exists some constant* $C_T<\infty$ *that only depends on the time parameter* $T$ *such that*

$$\sup_{0\leqslant t\leqslant T}\ \sup_{\varphi\in\mathscr{D}}\ \mathbb{E}(|\gamma_t^{(N)}(\varphi)-\gamma_t(\varphi)|)\leqslant\frac{C_T}{\sqrt{N}}. \tag{18}$$

**Proof.** To see this claim we plainly use the decomposition

$$\gamma_t^{(N)}(f)-\gamma_t(f)$$
$$=(\mathrm{e}^{\int_0^t \eta_s^{(N)}(U_s)\,\mathrm{d}s}-\mathrm{e}^{\int_0^t \eta_s(U_s)\,\mathrm{d}s})\ \eta_t^{(N)}(f)+(\eta_t^{(N)}(f)-\eta_t(f))\ \mathrm{e}^{\int_0^t \eta_s(U_s)\,\mathrm{d}s}.$$

From this and the fact that

$$|\mathrm{e}^{\int_0^t \eta_s^{(N)}(U_s)\,\mathrm{d}s}-\mathrm{e}^{\int_0^t \eta_s(U_s)\,\mathrm{d}s}|\leqslant 2\mathrm{e}^{T\|U\|_T^\star}\int_0^t |\eta_s^{(N)}(U_s)-\eta_s(U_s)|\,\mathrm{d}s\quad \forall t\leqslant T,$$

where

$$\|\ U\ \|_T^\star=\sup_{0\leqslant t\leqslant T}\|\ U_t\ \|,$$

it follows that for any $f\in\mathscr{D}$ and any $0\leqslant t\leqslant T$,

$$|\gamma_t^{(N)}(f)-\gamma_t(f)|\leqslant 2T\mathrm{e}^{T\|U\|_T^\star}(|\eta_t^{(N)}(f)-\eta_t(f)|+|\eta_t^{(N)}(U_t)-\eta_t(U_t)|)$$

and the proof of (18) is now a straighforward consequence of Theorem 4.2. □

**Remark 4.5.** From the density of $\mathscr{A}$ in $\mathscr{C}_b(E)$, it is easy to deduce that for any $t\geqslant 0$, $\eta_t^{(N)}$ weakly converges to $\eta_t$ in probability. That is, for any open neighbourhood $\mathscr{V}$ of $\eta_t$

$$\lim_{N\to\infty}\mathbb{P}(\eta_t^{(N)}\notin\mathscr{V})=0.$$

Furthermore, the inclusion $\mathscr{A}\subset\mathscr{C}_b(E)$ is not strictly necessary (the properties of martingales problems used here are satisfied in the $\mathscr{B}_b(E)$ context, cf. Ethier and Kurtz (1986), except that the càdlàgicity of the martingales appearing in (H1) has now to be assumed), and this fact will be applied in the case of bounded generators, see the first example of Section 5.1.3.

## 5. Example of mutations generators

We present here strong and general conditions implying the hypothesis (H3) put on the mutation pregenerators family $L$. But we will see that these conditions are sufficient to treat the more classical examples of Markov processes.

### 5.1. Time-homogeneous mutations

Condition (H3) is hard to work with in practice. In order to obtain a more tractable condition it is convenient to first examine the "time-homogeneous" situation. Let us assume that the pregenerators do not depend on time. We will note by $L$ this pregenerator, defined on a dense subalgebra $\mathscr{A} \subset \mathscr{C}_b(E)$, and $\Gamma$ its corresponding "carré du champ" bilinear operator. The semigroup will satisfy for all $0 \leqslant s \leqslant t$, $P_{s,t} = P_{0,t-s}$, and we will use the obvious notation $P_t = P_{0,t}$, for $t \geqslant 0$.

We also suppose that the fitness functions $U_t$, $t \geqslant 0$, do not depend on time, and naturally we note $U = U_0$.

Let us make the following hypothesis:

(H4) $U \in \mathscr{A}$.

In particular, under (H4) the functions $U$ and $\Gamma(U,U)$ are elements of $\mathscr{C}_b(E)$ and we can define

$$C^{(4)} = \sqrt{2(\| \Gamma(U,U) \| + \| U \|^2)}.$$

We also use the following assumption:

(H5) *For any $T$ there exists a constant $C_T^{(5)} < \infty$ such that for any $f \in \mathscr{A}$ and $t \in [0,T]$*

$$\| \Gamma(P_t f, P_t f) \| \leqslant C_T^{(5)} \max(\| f \|^2; \| \Gamma(f,f) \|).$$

We begin our program with:

**Proposition 5.1.** *Assume that* (H4) *and* (H5) *hold. Then* (H3) *is satisfied with the subset $\mathscr{D} \subset \mathscr{A}$ given by*

$$\mathscr{D} = \{ f \in \mathscr{A}: \ \max(\| f \|; \| \Gamma(f,f) \|) \leqslant 1 \}$$

*and with the constants $C_T^{(1)} = C^{(4)}$, $C_T^{(2)} = C_T^{(5)}$ and $C_T^{(3)} = C_T^{(5)}$.*

It then follows that we can take for jauge

$$\forall f \in \mathscr{A}, \qquad |||f||| = \max(\| f \|; \sqrt{\| \Gamma(f,f) \|}).$$

**Proof.** Let us first show that

$$\forall f \in \mathscr{D}, \forall t \in [0,T], \qquad P_t(f) \in \sqrt{C_T^{(5)}} \mathscr{D}.$$

For this we simply note that for any $f \in \mathscr{D}$ and $0 \leqslant t$

$$\| P_t(f) \| \leqslant \| f \| \leqslant 1$$

and, under (H5)

$$\| \Gamma(P_t(f),P_t(f)) \| \leqslant C_T^{(5)}.$$

To complete the proof of the proposition, it clearly suffices to check that

$$\forall f \in \mathscr{D}, \qquad Uf/C^{(4)} \in \mathscr{D}.$$

To see this claim, we note that for any $f \in \mathscr{D}$,

$$\| Uf \| \leqslant \| U \| \| f \| \leqslant \| U \| \leqslant C^{(4)}.$$

On the other hand, some elementary computations yield (where we use the general approximation $\Gamma(g,g) = \lim_{t\to 0_+} P_t((g-g(x))^2)(x)/t$, valid for all $g \in \mathscr{A}$),

$$\| \Gamma(Uf,Uf) \| \leqslant 2(\| f \|^2 \| \Gamma(U,U) \| + \| U \|^2 \| \Gamma(f,f) \|)$$

and therefore

$$\| \Gamma(Uf,Uf) \| \leqslant 2(\| \Gamma(U,U) \| + \| U \|^2) = (C^{(4)})^2. \qquad \square$$

Very often, the simplest way to verify hypothesis (H5), is to impose a lower bound on a curvature associated to the pregenerator $L$. We now make the assumption that $L(\mathscr{A}) \subset \mathscr{A}$, and then for $\alpha \geqslant 0$, let $R_\alpha \in \mathbb{R} \sqcup \{-\infty\}$ the largest constant such that for all $f \in \mathscr{A}$,

$$\Gamma(L(f),f) \leqslant -R_\alpha \Gamma(f,f) + \frac{\alpha}{2} L(\Gamma(f,f)) \tag{19}$$

(as usual, if there is no such finite constant, we put $R_\alpha = -\infty$), and we define the modified curvature constant as the "number"

$$R \stackrel{\text{def.}}{=} \sup_{\alpha \geqslant 0} R_\alpha \in \mathbb{R} \sqcup \{-\infty\}.$$

In the literature, the curvature associated to a pregenerator $L$ is given as $R_1$ (cf. Bakry (1994) for diffusion pregenerators and Schmuckenschläger (1998) for jumps pregenerators), because it is then the largest constant such that for all $f \in \mathscr{A}$,

$$\Gamma_2(f,f) \geqslant R_1 \Gamma(f,f),$$

where $\Gamma_2$ is naturally defined by

$$\forall f,g \in \mathscr{A}, \qquad \Gamma_2(f,g) = \tfrac{1}{2}(L(\Gamma(f,g)) - \Gamma(L(f),g) - \Gamma(f,L(g)))$$

and in case $L$ is the Laplacian on a Riemannian manifold, one gets for $R_1$ the usual Ricci curvature.

But in fact, it can be easily proved that in the situation where $L$ is the pregenerator of a non-degenerate diffusion on a manifold (for the definitions, see for instance Ikeda and Watanabe, 1981) and where $\mathscr{A}$ contains at least all smooth functions with compact support, one has $R_\alpha = -\infty$ if $\alpha \neq 1$ (because only when $\alpha = 1$, it is possible to have cancelations of some third-order derivatives of the functions), thus $R = R_1$. Nevertheless, in general, $R$ is a little better than $R_1$, for instance, it can be shown by direct calculations that for the asymmetric Bernoulli processes on two points, $R = R_0 > R_1$. So perhaps $R$ is a good definition of the curvature of a pregenerator $L$, however that may be, for our purposes, only $R$ will be needed.

**Proposition 5.2.** *Let us assume a strong regularity in time of the semi-group: for all* $t \geqslant 0$, *we suppose that in the* $|||\cdot|||$ *sense, on* $\mathscr{A}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}P_t = P_t L = L P_t.$$

*Then with the previous notations, one is always assured that*

$$\forall f \in \mathscr{A},\ \forall T \geqslant 0, \quad \Gamma(P_T(f), P_T(f)) \leqslant \exp(-2RT)\Gamma(f,f)$$

*so if we suppose that* $R > -\infty$, *then* (H5) *is satisfied with* $C_T^{(5)} = \exp((-2R)_+ T)$.

**Proof.** We will first consider the case $\alpha > 0$. Let $T > 0$ and $f \in \mathscr{A}$ be fixed, and define for $0 \leqslant t \leqslant T$,

$$F_t \stackrel{\text{def.}}{=} \Gamma(P_t(f), P_t(f)),$$

$$G_t \stackrel{\text{def.}}{=} P_{\alpha(T-t)}(F_t),$$

which are non-negative functions.

One is lead to differentiate $G_t$ in time (in the $\|\cdot\|$ sense), to get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}G_t &= -\alpha L(P_{\alpha(T-t)}(F_t)) + P_{\alpha(T-t)}\left(\frac{\mathrm{d}}{\mathrm{d}t}F_t\right) \\
&= -\alpha P_{\alpha(T-t)}(L(F_t)) + 2P_{\alpha(T-t)}(\Gamma(L(P_t(f)), P_t(f))) \\
&\leqslant -2R_\alpha P_{\alpha(T-t)}(F_t) \\
&= -2R_\alpha G_t,
\end{aligned}$$

where we have applied (19) with $f$ replaced by $P_t(f) \in \mathscr{A}$. This differential inequality is integrated at once to provide the upper bound

$$F_T = G_T \leqslant \exp(-2R_\alpha T)G_0 = \exp(-2R_\alpha T)P_{\alpha T}(F_0)$$

from which it follows that

$$\| F_T \| \leqslant \exp(-2R_\alpha T) \| F_0 \| . \tag{20}$$

This is also true for $\alpha = 0$, since it is then enough to differentiate directly $F_t$ in time. So as (20) is always satisfied for $\alpha \geqslant 0$, one has

$$\| F_T \| \leqslant \exp(-2RT) \| F_0 \|$$

and the proposition is proved. □

### 5.2. Time-inhomogeneous setting

We follow the ideas introduced below for homogeneous Markov processes.

So let us assume a strong (but quite usual in the context of semigroup approaches in probability) regularity of the inhomogeneous semigroup $(P_{s,t})_{0\leqslant s\leqslant t}$: for all $0 \leqslant s \leqslant t$, we suppose that in the $|||\cdot|||_0$ sense,

$$\frac{d}{ds}P_{s,t} = -L_s P_{s,t},$$

$$\frac{d}{dt}P_{s,t} = P_{s,t}L_t, \tag{21}$$

where $|||\cdot|||_0$ is the norm given on $\mathscr{A}$ by

$$\forall f \in \mathscr{A}, \qquad |||f|||_0 = \max(\| f \|; \sqrt{\| \Gamma_0(f,f) \|}).$$

We will need even a little more: let $\alpha:\mathbb{R}_+ \to \mathbb{R}_+^*$ a measurable and integrable function be given, we make the assumption that there exists an inhomogeneous semigroup of Markov kernels $(P_{s,t}^{(\alpha)})_{0\leqslant s\leqslant t}$ such that,

$$\forall 0\leqslant s\leqslant t,$$

$$P_{s,s} = \mathrm{Id},$$

$$\frac{d}{dt}P_{s,t}^{(\alpha)} = \alpha_t P_{s,t}^{(\alpha)} L_t \tag{22}$$

(in the pointwise sense will be enough here, so for instance one can be lead to consider the martingales problem associated to the family of pregenerators $(\alpha_t L_t)_{t\geqslant 0}$).

Furthermore, we suppose that for all $f\in\mathscr{A}$, $\mathbb{R}_+ \ni t \mapsto \Gamma_t(f,f)$ is derivable (as before, in $\mathscr{C}_b(E)$), and we denote for $t\geqslant 0$,

$$A_t = \sup_{f\in\mathscr{A}} \left\| \frac{\partial_t \Gamma_t(f,f)}{\Gamma_t(f,f)} \right\|$$

(where $\partial_t$ stands for $d/dt$).

We also assume that for all $t\geqslant 0$, $L_t(\mathscr{A})\subset\mathscr{A}$, and that for all $f\in\mathscr{A}$, $\mathbb{R}_+\times E \ni (t,x) \mapsto L_t(L_t(f))(x)$ is measurable, this stability also enables us to consider $R_t$ the curvature of $L_t$, as it is defined in the previous subsection.

We suppose that both the applications $\mathbb{R}_+ \ni t \mapsto A_t$ and $\mathbb{R}_+ \ni t\mapsto R_t$ are locally in the $\mathbb{L}^1$-space for the Lebesgue measure.

Furthermore, for the latter application, we will assume that if we denote for $\alpha, t > 0$,

$$R_{\alpha,t} = \inf_{f\in\mathscr{A}, x\in E} \frac{\alpha L_t(\Gamma_t(f,f))(x) - \Gamma_t(L_t(f),f)(x)}{\Gamma_t(f,f)(x)},$$

then there exists a sequence $(\alpha_n)_{n\in\mathbb{N}}$ of measurable and integrable applications from $\mathbb{R}_+$ to $\mathbb{R}$, such that for all $T>0$,

$$\lim_{n\to\infty} \int_0^T R_{\alpha_n(s),s}\, ds = \int_0^T R_s\, ds,$$

(note that if the $L_t$, $t\geqslant 0$, are diffusion pregenerators, then this extra condition is automatically satisfied, as one can take $\alpha_n \equiv 1$, there is even no need for condition (22), since we will use it only for the applications $\alpha_n$, $n\geqslant 0$).

Finally, let us consider the constant

$$C_T^{(4)} = \sup_{0\leqslant t\leqslant T} \sqrt{2(\| \Gamma_0(U_t,U_t) \| + \| U_t \|^2)}$$

and suppose it is finite.

Under all the above hypotheses, we can see that

**Proposition 5.3.** *If we define*

$$\mathscr{D} = \{f \in \mathscr{A}: ||| f |||_0 \leqslant 1\},$$

*then for all $T > 0$ fixed, condition* (H3) *is fulfilled with*

$$C_T^{(1)} = C_T^{(4)},$$

$$C_T^{(2)} = \exp\left( \int_0^T 3A_t - 2R_t \, dt \right),$$

$$C_T^{(3)} = \exp\left( \int_0^T 2A_t - 2R_t \, dt \right).$$

Here the associated jauge is clearly $||| \cdot |||_0$.

**Proof.** The membership of $U_t\varphi$ to $C_T^{(4)}\mathscr{D}$, for $\varphi \in \mathscr{D}$, is shown as in the homogeneous case. Then we remark that for all $0 \leqslant s \leqslant t \leqslant T$ and all $\varphi \in \mathscr{A}$,

$$\exp\left( -\int_s^t A_u \, du \right) \Gamma_s(\varphi,\varphi) \leqslant \Gamma_t(\varphi,\varphi) \leqslant \exp\left( \int_s^t A_u \, du \right) . \Gamma_s(\varphi,\varphi)$$

So in fact, derivations (21) are also true in the sense of the norm $||| \cdot |||_u = \max(\| \cdot \|; \sqrt{\| \Gamma_u(\cdot,\cdot) \|})$, for all $u \in [0,T]$. A consequence of this property is that in the $\| \cdot \|$ sense, for all $\varphi \in \mathscr{A}$, the application $[0,T] \ni t \mapsto \Gamma_t(P_{t,T}(\varphi), P_{t,T}(\varphi))$ is derivable, and its derivative is

$$\frac{d}{dt}\Gamma_t(P_{t,T}(\varphi), P_{t,T}(\varphi)) = (\partial_t \Gamma_t)(P_{t,T}(\varphi), P_{t,T}(\varphi)) - 2\Gamma_t(L_t(P_{t,T}(\varphi)), P_{t,T}(\varphi)).$$

So let $s \geqslant 0$, $\varphi \in \mathscr{D}$ and $n \in \mathbb{N}$ be given, we consider for $s \leqslant t \leqslant T$,

$$G_t = P_{s,t}^{(\alpha_n)}[\Gamma_t(P_{t,T}(\varphi), P_{t,T}(\varphi))].$$

As before, we calculate that

$$\begin{aligned} \partial_t G_t &= \alpha_n(t) P_{s,t}^{(\alpha_n)}[L_t(\Gamma_t(P_{t,T}(\varphi), P_{t,T}(\varphi))] + P_{s,t}^{(\alpha_n)}[(\partial_t \Gamma_t)(P_{t,T}(\varphi), P_{t,T}(\varphi))] \\ &\quad - 2P_{s,t}^{(\alpha_n)}[\Gamma_t(L_t P_{t,T}(\varphi), P_{t,T}(\varphi))] \\ &\geqslant (2R_{\alpha_n(t),t} - A_t) G_t. \end{aligned}$$

By integrating this inequality, we get

$$G_T \geqslant G_s \exp\left( \int_s^T 2R_{\alpha_n(t),t} - A_t \, dt \right)$$

from where it follows that

$$\begin{aligned} \| \Gamma_s(P_{s,T}(\varphi), P_{s,T}(\varphi)) \| &\leqslant \| \Gamma_T(\varphi,\varphi) \| \exp\left( \int_s^T A_t - 2R_{\alpha_n(t),t} \, dt \right) \\ &\leqslant \| \Gamma_0(\varphi,\varphi) \| \exp\left( \int_0^T 2A_t - 2R_{\alpha_n(t),t}. \, dt \right). \end{aligned}$$

But letting $n$ go to infinity, we obtain that the two last requirements of condition (H3) are fulfilled with the constants presented in the proposition. □

### 5.3. Examples of mutations pregenerators

Here are some classical examples that can be handled in our framework.

#### 5.3.1. Bounded generators

The simplest examples of mutation pregenerators which satisfy the previous hypotheses are those of measurable families of bounded generators $(L_t)_{t\geqslant 0}$. Namely, for all $t\geqslant 0$, let $L_t : E\times\mathscr{E}\to\mathbb{R}$ be a signed kernel such that

- for any $(t,x)\in\mathbb{R}_+\times E$, $L_t(x,\cdot\cap(E\backslash\{x\}))\in\boldsymbol{M}(E)$ and $L_t(x,E)=0$,
- for any $A\in\mathscr{E}$, $\mathbb{R}_+\times E\ni(t,x)\mapsto L_t(x,A)\in\mathbb{R}_+$ is a measurable function,
- for all $T>0$, there exists a constant $0\leqslant M_T<\infty$ such that

$$\forall(t,x)\in[0,T]\times E,\qquad L_t(x,E\backslash\{x\})\leqslant M_T.$$

We can take here $\mathscr{A}=\mathscr{B}_{\mathrm{b}}(E)$, and in this case the inhomogeneous variant of (H5) clearly holds, because we have

$$\forall f\in\mathscr{A},\qquad \|\,\Gamma_t(f,f)\,\| = \sup_{x\in E}\frac{1}{2}\int(f(y)-f(x))^2L_t(x,\mathrm{d}y)$$
$$\leqslant 2M_T\,\|\,f\,\|\,.$$

So in this particular situation we can choose for the subset $\mathscr{D}$ the set of all measurable and bounded functions $f$ such that $\|\,f\,\|\leqslant 1$. In view of Theorem 4.2 for any $T>0$ there exists a constant $C_T<\infty$ such that

$$\sup_{0\leqslant t\leqslant T}\mathbb{E}(|\eta_t^N(f)-\eta_t(f)|)\leqslant\frac{C_T}{\sqrt{N}}\,\|\,f\,\|$$

for any measurable and bounded function $f : E\to\mathbb{R}$.

#### 5.3.2. Riemannian diffusions

Let $E$ be a compact Riemannian manifold. As usual, $\langle\cdot,\cdot\rangle$, $\nabla\cdot$ and $\triangle\cdot$ will denote the scalar product, the gradient (or more generally the connexion) and the Laplacian associated to this structure. Let $\mathscr{A}$ be the algebra of smooth functions, i.e. $\mathscr{A}=C^\infty(E)$. Suppose we are given a family $(b_t)_{t\in\mathbb{R}_+}$ of vector fields, such that

$$\mathbb{R}_+\times E\ni(t,x)\mapsto b_t(x)\in\boldsymbol{T}_x(E)\subset\boldsymbol{T}(E)$$

is smooth.

We denote for $t\geqslant 0$,

$$L_t : \mathscr{A}\to\mathscr{A}$$
$$f\mapsto\frac{\triangle f}{2}+\langle b_t,\nabla f\rangle.$$

It is immediate to check that in this example the carré du champ does not depend on time and is given for all $t \geqslant 0$, by

$$\forall f, g \in \mathscr{A}, \qquad \Gamma_t(f, f) = \langle \nabla f, \nabla g \rangle.$$

Classical results show that (H1) is fulfilled and that for all $T > 0$ and all $\varphi \in \mathscr{A}$, the application $F_{T,\varphi}$ is smooth in $[0, T] \times E$ (for instance, as a solution of a regular parabolic equation), from where it follows that (H2) is satisfied, since it is elementary to prove that $A(T, \mathscr{A})$ contains all smooth functions.

The strong assumption on the time regularity of the semigroup presented in Section 5.1.2 is also well known. Furthermore, as mentioned before, since the $L_t$, $t \geqslant 0$, are diffusion pregenerators, we have that

$$R_t = R_{1,t}$$

for which we have the lower bound (cf. for instance Section 6 of Bakry, 1994)

$$R_t \geqslant \frac{R}{2} - \sup_{x \in E} r_t(x),$$

where $R$ is the Ricci curvature of $E$ and where $r_t(x)$ is the largest eigenvalue of $\nabla^{\mathrm{s}} b_t(x)$, which is the symmetrized tensor associated to the tensor $\nabla b_t(x)$. More precisely, in local coordinates, $\nabla^{\mathrm{s}} b_t(x)$ is given by the symmetrization of the matrix

$$\left( \sum_{1 \leqslant l \leqslant d} g^{i,l}(x) \left[ \partial_l b_t^j(x) + \sum_{1 \leqslant k \leqslant d} \Gamma^j_{l,k}(x) b_t^k(x) \right] \right)_{1 \leqslant i,j \leqslant d},$$

where $d$ is the dimension of $E$, $(g^{i,j}(x))_{1 \leqslant i,j \leqslant d}$ is the inverse of the matrix defining the scalar product in $\boldsymbol{T}_x(E)$, and $(\Gamma^i_{j,l}(x))_{1 \leqslant i,j,l \leqslant d}$ are the Christoffel symbols in the point $x \in E$ of the connexion $\nabla$.

It then easily follows that assumption (H3) is also verified, via Proposition 5.3.

#### *5.3.3. Euclidean diffusions*

Except for the compacity of the state space, these processes are similar to those of the previous example.

So here $E = \mathbb{R}^d$, $d \geqslant 1$, and let for $(t, x) \in \mathbb{R}_+ \times E$, $a(t, x) = (a^{i,j}(t, x))_{1 \leqslant i,j \leqslant d}$ be a symmetric positive-definite matrix. We suppose they are uniformly elliptic: there exists a constant $\varepsilon > 0$ such that for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$,

$$\forall y = (y_i)_{1 \leqslant i \leqslant d} \in \mathbb{R}^d, \qquad \sum_{1 \leqslant i,j \leqslant d} a^{i,j}(t, x) y_i y_j \geqslant \varepsilon \sum_{1 \leqslant i \leqslant d} y_i^2.$$

We, furthermore, assume that the applications $a^{i,j} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$, $1 \leqslant i, j \leqslant d$, belong to $C_{\mathrm{b}}^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$.

Let also $b = (b^i)_{1 \leqslant i \leqslant d} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ be a $C_{\mathrm{b}}^{0,1}(\mathbb{R}_+ \times \mathbb{R}^d)$ application.

We denote $\mathscr{A} = C_{\mathrm{b}}^2(\mathbb{R}^d)$ and we consider on it the generators $L_t$, $t \geqslant 0$, given by

$$\forall f \in \mathscr{A}, \forall x \in \mathbb{R}^d, \quad L_t(f)(x) = \sum_{1 \leqslant i,j \leqslant d} \frac{a^{i,j}}{2}(t, x) \partial_{i,j} f(x) + \sum_{1 \leqslant i \leqslant d} b^i(t, x) \partial_i f(x).$$

For $(t,x)\in\mathbb{R}_+\times E$, it is well known that (H1) is fulfilled (for more details about this problem, see Stroock and Varadhan (1979)) and that $\mathbb{P}_{t,x}$ is the law of the solution of the stochastic differential equation

$$X_t = x,$$

$$\mathrm{d}X_s = \sigma(s,X_s)\,\mathrm{d}B_s + b(s,X_s)\,\mathrm{d}s, \quad s\geqslant t,$$

where $\sigma$ is an application from $\mathbb{R}_+\times\mathbb{R}^d$ into the set of symmetric definite positive matrices such that for all $(t,x)\in\mathbb{R}_+\times\mathbb{R}^d$, $\sigma(t,x)^2 = a(t,x)$, and where $(B_t)_{t\geqslant 0}$ is a standard $d$-vector Brownian motion.

But as before, (H2) and (H3) are also classical results. In fact, for $t\geqslant 0$ given, one can take

$$A_t = \frac{\left\|\sqrt{\sum_{1\leqslant i,j\leqslant d}(\partial_t a_{i,j}(t,\cdot))^2}\right\|}{\varepsilon}$$

and one can find a lower bound of $R_t$ in term of $\varepsilon$, $\max_{1\leqslant i,j,k,l\leqslant d}\|\partial_{k,l}a^{i,j}(t,\cdot)\|$, $\max_{1\leqslant i,j,k\leqslant d}\|\partial_k a^{i,j}(t,\cdot)\|$, $\max_{1\leqslant i,j\leqslant d}\|a^{i,j}(t,\cdot)\|$, $\max_{1\leqslant i,j\leqslant d}\|\partial_j b^i(t,\cdot)\|$ and $\max_{1\leqslant i\leqslant d}\|b^i(t,\cdot)\|$ (cf. for instance Bakry, 1994 or Ikeda and Watanabe, 1981).

The situation is particularly simple when the diffusion matrices are constant, and let us consider the case where $(a^{i,j})_{1\leqslant i,j\leqslant d}\equiv \mathrm{Id}$. Then we have

$$R_t = -\sup_{x\in\mathbb{R}^d} r_t(x),$$

where for $x\in\mathbb{R}^d$, $r_t(x)$ is the largest eigenvalue of the symmetric matrix

$$\tfrac{1}{2}\left(\partial_i b^j(t,x) + \partial_j b^i(t,x)\right)_{1\leqslant i,j\leqslant d},$$

so we are assured of

$$R_t \geqslant -d \max_{1\leqslant i,j\leqslant d}\|\partial_j b^i(t,\cdot)\|.$$

## Acknowledgements

We heartily thank Carl Graham, Jean Jacod and Sylvie Méléard for encouraging and fruitful discussions and the anonymous referee for his comments and suggestions that greatly improved the presentation of the manuscript.

## References

Bakry, D., 1994. L'hypercontractivité et son utilisation en théorie des semigroupes. In: Bernard, P. (Ed.), Lectures on Probability Theory, Ecole d'Eté de Probabilités de Saint-Flour XXII-1992, Lecture Notes in Mathematics, Vol. 1581. Springer, Berlin.

Dawson, D., 1993. Measure–valued Markov processes. In: Hennequin, P.L. (Ed.), Ecole d'Eté de Probabilités de Saint–Flour XXI–1991, Lecture Notes in Mathematics, Vol. 1541. Springer, Berlin.

Ethier, S., Kurtz, T., 1986. Markov Processes, Characterization and Convergence, Wiley Series in Probability and Mathematical Statistics. Wiley, New York.

Graham, C., Méléard, S., 1997. Stochastic particle approximations for generalized Boltzmann models and convergence estimates. Ann. Probab. 25 (1), 115–132.

Ikeda, N., Watanabe, S., 1981. Stochastic Differential Equations and Diffusion Processes. North-Holland Mathematical Library, Vol. 24. North-Holland/Kodansha, Amsterdam.

Krylov, N.V., 1964. Partial Differential Equations of Parabolic Type. Prentice-Hall, Englewood Cliff, New York.

Liggett, T.M., 1985. Interacting Particle Systems, A Series of Comprehensive Studies in Mathematics, Vol. 276. Springer, New York.

Méléard, S., 1996. Asymptotic behaviour of some interacting particle systems, McKean–Vlasov and Boltzmann models. In: Talay, D., Tubaro, L. (Eds.), Probabilistic Models for Nonlinear Partial Differential Equations, Montecatini Terme, 1995, Lecture Notes in Mathematics, Vol. 1627. Springer, Berlin.

Schmuckenschläger, M., 1998. Curvature of non-local Markov generators. Préprint.

Stroock, D., Varadhan, S.R.S., 1979. Multidimensional Diffusion Processes. Springer, Berlin.

Sznitman, A.S., 1991. Topics in propagation of chaos. In: Hennequin, P.L. (Ed.), Ecole d'Eté de Probabilités de Saint–Flour XIX-1989, Lecture Notes in Mathematics, Vol. 1464. Springer, Berlin.

Sznitman, A.S., 1997. Brownian Motion, Obstacles and Random Media, Springer Monographs in Mathematics, Springer, 1998.