



**HAL**  
open science

# Invariant measure of duplicated diffusions and application to Richardson-Romberg extrapolation

Vincent Lemaire, Gilles Pagès, Fabien Panloup

► **To cite this version:**

Vincent Lemaire, Gilles Pagès, Fabien Panloup. Invariant measure of duplicated diffusions and application to Richardson-Romberg extrapolation. 2013. hal-00785766v1

**HAL Id: hal-00785766**

**<https://hal.science/hal-00785766v1>**

Preprint submitted on 6 Feb 2013 (v1), last revised 8 Jul 2014 (v5)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Invariant measure of duplicated diffusions and application to Richardson-Romberg extrapolation

Vincent Lemaire\*, Gilles Pagès†, Fabien Panloup‡

February 6, 2013

## Abstract

With a view to numerical applications we address the following question: given an ergodic Brownian diffusion with an unique invariant measure, what are the invariant measures of the duplicated system consisting of two trajectories ? We focus on the interesting case where the two trajectories follow the same Brownian path, and give explicit conditions on the drift and diffusion coefficient function to obtain uniqueness for invariant distribution of the duplicated system.

As an application, we investigate the Richardson-Romberg extrapolation for the numerical approximation of the invariant measure of the initial ergodic Brownian diffusion.

*Keywords:* Invariant measure; ergodic diffusion; Central Limit Theorem; Euler scheme; Richardson-Romberg extrapolation.

*AMS classification (2000):* 60G10, 60J60, 65C05, 60F05.

## 1 Introduction and motivations

When one discretizes a (stochastic or not) differential equation by an Euler scheme with step  $h$ , a classical method to reduce the discretization error is the so-called Richardson-Romberg (*RR*) extrapolation. Roughly speaking, the idea of this method is to introduce a second Euler scheme with step  $h/2$  and to choose an appropriate linear combination of the two schemes to cancel the first-order discretization error. Such an idea can be adapted to the long-time setting. More precisely, when one tries to approximate the invariant distribution of a diffusion by empirical measures based on an Euler scheme (with decreasing step) of the diffusion, it is also possible to implement the same strategy by introducing a second Euler scheme with half-step. In fact, tackling the rate of convergence of such a procedure involving a couple of Euler schemes of the same diffusion leads to studying the long run behaviour of the underlying couple of continuous processes that we will call *duplicated diffusion* (also known as *2-point motion* in some special setting as seen further

---

\*Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, UPMC, Case 188, 4 pl. Jussieu, F-75252 Paris Cedex 5, France, E-mail: [vincent.lemaire@upmc.fr](mailto:vincent.lemaire@upmc.fr)

†Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, UPMC, Case 188, 4 pl. Jussieu, F-75252 Paris Cedex 5, France, E-mail: [gilles.pages@upmc.fr](mailto:gilles.pages@upmc.fr)

‡Institut de Mathématiques de Toulouse, Université Paul Sabatier & INSA Toulouse, 135, av. de Rangueil, F-31077 Toulouse Cedex 4, France, E-mail: [fabien.panloup@math.univ-toulouse.fr](mailto:fabien.panloup@math.univ-toulouse.fr)

on). Before being more specific as concerns this motivation, let us now define what a duplicated diffusion is.

Consider the following Brownian diffusion solution to the stochastic differential equation

$$(SDE) \quad \equiv \quad dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d, \quad (1.1)$$

where  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}(d, q, \mathbb{R})$  ( $d \times q$  matrices with real valued entries) are locally Lipschitz continuous with linear growth and  $W$  is a standard  $q$ -dimensional Brownian motion defined on a filtered probability space  $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$  (satisfying the usual conditions). This stochastic differential equation (SDE) has a unique strong solution denoted  $X^x = (X_t^x)_{t \geq 0}$ . Let  $\rho \in \mathcal{M}(q, q, \mathbb{R})$  be a square matrix with transpose  $\rho^*$  such that  $I_q - \rho\rho^*$  is non-negative as a symmetric matrix. We consider a filtered probability space, still denoted  $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$  on which is defined a  $2q$ -dimensional standard  $(\mathcal{F}_t)$ -Brownian motion denoted  $(W, \widetilde{W})$  so that  $W$  and  $\widetilde{W}$  are two *independent*  $q$ -dimensional standard  $(\mathcal{F}_t)$ -Brownian motions. Then we define  $W^{(\rho)}$  a third standard  $q$ -dimensional  $(\mathcal{F}_t)$ -Brownian motions by

$$W^{(\rho)} = \rho^*W + \sqrt{I_q - \rho^*\rho}\widetilde{W},$$

which clearly satisfies

$$\langle W^i, W^{(\rho),j} \rangle_t = \rho_{ij} t, \quad t \geq 0$$

(the square root should be understood in the set of symmetric non-negative matrices). The duplicated diffusion or “duplicated stochastic differential system” (DSDS) is then defined by

$$(DSDS) \quad \equiv \quad \begin{cases} dX_t = b(X_t)dt + \sigma(X_t)dW_t, & X_0 = x_1 \in \mathbb{R}^d, \\ dX_t^{(\rho)} = b(X_t^{(\rho)})dt + \sigma(X_t^{(\rho)})dW_t^{(\rho)}, & X_0^{(\rho)} = x_2 \in \mathbb{R}^d. \end{cases} \quad (1.2)$$

Under the previous assumptions on  $b$  and  $\sigma$ , (1.2) has a unique (strong) solution. Then both  $(X_t^x)_{t \geq 0}$  and  $(X_t^{x_1}, X_t^{(\rho),x_2})_{t \geq 0}$  are homogeneous Markov processes with transition (Feller) semi-groups, denoted  $(P_t(x, dy))_{t \geq 0}$  and  $(Q_t^{(\rho)}((x_1, x_2), (dy_1, dy_2)))_{t \geq 0}$  respectively, and defined on test Borel functions  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $g : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ , by

$$P_t(f)(x) = \mathbb{E}f(X_t^x) \quad \text{and} \quad Q_t^{(\rho)}(g)(x, x') = \mathbb{E}g(X_t^{x_1}, X_t^{(\rho),x_2}).$$

We will assume throughout the paper that the original diffusion  $X^x$  has an unique invariant distribution denoted  $\nu$  *i.e.* satisfying  $\nu P_t = \nu$  for every  $t \in \mathbb{R}_+$ . The first aim of this paper is to elucidate what are the invariant distributions of  $(Q_t^{(\rho)})_{t \geq 0}$  (if any). Thus, if  $\rho = 0$ , it is clear that  $\nu \otimes \nu$  is invariant for  $Q^{(0)}$  and if  $\rho = I_q$  so is  $\nu_\Delta = \nu \circ (x \mapsto (x, x))^{-1}$ , but are they the only ones? In the sequel, we will denote by  $\mu$  a generic invariant measure of  $Q^{(\rho)}$ . In view of the application to the *RR* extrapolation method for the computation by simulation of the invariant measure  $\nu$  that will be analyzed in Section 4, this question of the (existence and) uniqueness of  $\mu$  is crucial to justify and support this method.

▷ *Existence of an invariant distribution for  $(Q_t^{(\rho)})_{t \geq 0}$ .* First, the family of probability measures  $(\mu_t^{(\rho)})_{t > 0}$  defined on  $(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d)^{\otimes 2})$  by

$$\mu_t^{(\rho)} = \frac{1}{t} \int_0^t \nu^{\otimes 2}(dx_1, dx_2) Q_s^{(\rho)}((x_1, x_2), (dy_1, dy_2)) ds$$

is tight since both its marginals on  $\mathbb{R}^d$  are equal to  $\nu$ . Furthermore, the semi-group  $(Q_t^{(\rho)})_{t \geq 0}$  being Feller, one easily shows that any of its limiting distributions  $\mu^{(\rho)}$  as  $t \rightarrow \infty$  is an invariant distribution for  $(Q_t^{(\rho)})_{t \geq 0}$  such that  $\mu^{(\rho)}(dx \times \mathbb{R}^d) = \mu^{(\rho)}(\mathbb{R}^d \times dx) = \nu(dx)$ . Also note that, if uniqueness fails and  $(P_t)_{t \geq 0}$  has two distinct invariant distributions  $\nu$  and  $\nu'$ , a straightforward adaptation of the above (sketch of) proof shows that  $(Q_t^{(\rho)})_{t \geq 0}$  has (at least) an invariant distribution with marginals  $(\nu, \nu')$  and another with  $(\nu', \nu)$  as marginals.

▷ *Uniqueness of the invariant distribution of  $(Q_t^{(\rho)})_{t \geq 0}$ .* It is clear that in full generality the couple  $(X, X^{(\rho)})$  may admit several invariant distributions even if  $X$  has only one such distribution. So is the case of a fully degenerate setting ( $\sigma \equiv 0$ ) if the flow  $\Phi(x, t)$  of the *ODE*  $\equiv \dot{x} = b(x)$  has 0 as a unique repulsive equilibrium and a unique invariant distribution  $\nu$  on  $\mathbb{R}^d \setminus \{0\}$ . Then both distributions  $\nu^{\otimes 2}$  and  $\nu_\Delta$  (defined as above) on  $(\mathbb{R}^d \setminus \{0\})^2$  are invariant and if  $\nu$  is not reduced to a Dirac mass (think *e.g.* to a 2-dimensional ODE with a limit cycle around 0) (*DSDS*) has at least two invariant distribution.

In the non-degenerate case ( $\sigma \neq 0$ ) the situation is more involved and depends on the correlation structure  $\rho$  between the two Brownian motions  $W$  and  $W^{(\rho)}$ . The diffusion matrix  $\Sigma(X_t^{x_1}, X_t^{(\rho), x_2})$  of the couple  $(X^{x_1}, X^{(\rho), x_2})$  at time  $t > 0$  is given by any continuous solution to the equation

$$\Sigma(\xi_1, \xi_2)\Sigma(\xi_1, \xi_2)^* = \begin{bmatrix} \sigma\sigma^*(\xi_1) & \sigma(\xi_1)\rho\sigma^*(\xi_2) \\ \sigma(\xi_2)\rho^*\sigma^*(\xi_1) & \sigma\sigma^*(\xi_2) \end{bmatrix}$$

(*e.g.* the square root in the symmetric non-negative matrices or the Choleski transform. . .).

First, note that if  $I_q - \rho^*\rho$  is positive definite as a symmetric matrix, it is straightforward that ellipticity or uniform ellipticity of  $\sigma\sigma^*$  (when  $q \geq d$ ) for  $X^x$  is transferred to  $\Sigma(X_t^{x_1}, X_t^{(\rho), x_2})\Sigma(X_t^{x_1}, X_t^{(\rho), x_2})^*$  for the couple  $(X^{x_1}, X^{(\rho), x_2})$ . Now, uniform ellipticity, combined with standard regularity and growth/boundedness assumption on the coefficients  $b$ ,  $\sigma$  and their partial derivatives, classically implies the existence for every  $t > 0$  of a (strictly) positive probability density  $p_t(x, y)$  for  $X_t^x$ . These additional conditions are automatically satisfied by the “duplicated coefficients” of (*DSDS*). At this stage, it is classical background that any homogeneous Markov process whose transition has a positive density for every  $t > 0$  has at most one invariant distribution (if any). Consequently, under these standard assumptions on  $b$  and  $\sigma$  which ensure uniqueness of the invariant distribution  $\nu$  for  $X$ , we get uniqueness for the “duplicated” diffusion process  $(X, X^{(\rho)})$  as well. The *hypo-elliptic case* which also implies the existence of a density for  $X_t^x$  and the uniqueness of the invariant distribution under *controllability assumptions* on a companion differential system of an SDE can also be transferred to (*DSDS*), although the proof becomes significantly less straightforward (see Appendix B for a precise statement and a detailed proof).

The second setting of interest,  $\rho = I_q$ , corresponds to  $X^{(\rho), x_2} = X^{x_2}$  (since  $W^{(\rho)} = W$ ). Note that in this very case the duplicated system is also known as the two-point motion associated to the Brownian flow (see [Kun90] section 4.2, [Har81]). In our application to the *RR* extrapolation for the recursive computation of the invariant distribution this setting amounts to the simulation of consistent Brownian increments for the two Euler schemes of (*SDE*) under consideration. Furthermore we will see that such a choice shares some optimality properties in terms of variance reduction. However, from an ergodic viewpoint the situation becomes dramatically different since the resulting duplicated stochastic

system becomes completely degenerate even if the original diffusion is uniformly elliptic with smooth coefficients. It is clear that

$$\nu_\Delta = \nu \circ (x \mapsto (x, x))^{-1}$$

is an invariant distribution for  $(DSDS)$  of  $\nu$  is for  $(SDE)$ , but the question becomes: “is it the only one?” (when  $\nu$  is unique).

We will show in the sequel of the paper that the answer is essentially positive in the one-dimensional case and negative in general in the multidimensional case. An additional *weak asymptotic confluence* condition is needed to ensure the uniqueness of  $\nu_\Delta$  and counterexamples can be exhibited when it fails.

The paper is organized as follows: in Section 2, we establish a one-dimensional result which grants the uniqueness of  $\nu_\Delta$  as an invariant distribution for  $(DSDS)$ . We also prove that the paths of the  $(DSDS)$  are asymptotically confluent (slightly extending by a different method a result by Has’minskii in [Has80]). Section 3, is devoted to the multidimensional case. We first provide a simple counter-example to uniqueness and then provide rather general but operating criterions, all based on the behavior of the one-dimensional Itô process  $|X_t^{x_1} - X_t^{x_2}|$  and closely related in spirit to the Feller classification for one-dimensional diffusions. They can be seen as asymptotic weak confluence criterions.

Section 4 is devoted to an application, namely to the recursive computation of the invariant measure of a diffusion by weighted occupation measures of an Euler scheme with decreasing step: we show how to improve the (weak) rate of convergence of the procedure by implementing a Richardson-Romberg  $(RR)$  extrapolation involving two Euler schemes with different steps. In particular, we will prove in this section that generating the two Euler schemes from the same underlying Brownian motion is the optimal way to proceed (in terms of variance reduction) if the duplicated system  $(DSDS)$  has  $\nu_\Delta$  as a unique invariant distribution.

NOTATIONS.  $\bullet \|\cdot\|$  denotes the canonical Euclidean norm on  $\mathbb{R}^d$ .

- $\|A\| = \sqrt{\text{Tr}(AA^*)}$  if  $A \in \mathcal{M}(d, q, \mathbb{R})$  (which is but the canonical Euclidean norm on  $\mathbb{R}^{d^2}$ ).
- $\Delta_{\mathbb{R}^{2d}} = \{(x, x), x \in \mathbb{R}^d\}$  denotes the diagonal of  $\mathbb{R}^{2d}$ .
- $\mathcal{S}(d, \mathbb{R}) = \{S \in \mathcal{M}(d, d, \mathbb{R}), S^* = S\}$  and  $\mathcal{S}^+(d, \mathbb{R})$  the subset of  $\mathcal{S}(d, \mathbb{R})$  of non-negative matrices.
- $\mu_n \xrightarrow{(\mathbb{R}^d)} \mu$  denotes the weak convergence of the sequence  $(\mu_n)_{n \geq 1}$  of probability measures defined on  $(\mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d))$  toward the probability measure  $\mu$ .
- For every function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , define the Lipschitz coefficient of  $f$  by  $[f]_{\text{Lip}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|} \leq +\infty$ .

## 2 The one-dimensional case

We first show that, in the one-dimensional case  $d = q = 1$ , uniqueness of  $\nu$  implies that  $\nu_\Delta$ , as defined in the introduction, is the unique invariant distribution of the duplicated diffusion. The main theorem of this section is Theorem 2.1 which consists of two claims. The first one establishes this uniqueness claim using some ergodic-type arguments. Note that we do not require that  $\sigma$  never vanishes. The second claim is an asymptotic pathwise confluence property for the diffusion in its own scale, established under some slightly more stringent assumptions involving the scale function  $p$ , see below. This second result, under

slightly less general assumptions, is originally due to Has'minskii (see [Has80], Appendix to the English edition, Theorem 2.2, p.308). It is revisited here by different techniques, mainly comparison results for one dimensional diffusions and ergodic arguments. Note that uniqueness of  $\nu_\Delta$  can always be retrieved from asymptotic confluence (see Remark 2.1).

Before stating the result, let us recall some definitions. We denote by  $M$  the *speed measure* of the diffusion classically defined by  $M(d\xi) = \frac{d\xi}{(\sigma^2 p')(\xi)}$ , where  $p$  is the *scale function* defined (up to a constant) by

$$p(x) = \int_{x_0}^x d\xi e^{-\int_{x_0}^\xi \frac{2b}{\sigma^2}(u)du}, \quad x \in \mathbb{R}.$$

Obviously, we will consider  $p$  only when it makes sense *as a finite function* (so is the case if  $b/\sigma^2$  is locally integrable on the real line). We are in position to state the result.

**THEOREM 2.1.** *Assume that  $b$  and that  $\sigma$  are continuous functions on  $\mathbb{R}$  being such that strong existence, pathwise uniqueness and the Feller Markov property hold for (SDE) from any  $x \in \mathbb{R}$ . Assume furthermore that there exists  $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , strictly increasing, with  $\lambda(0) = 0$  and  $\int_{0^+} \lambda(u)^{-2} du = +\infty$  such that for all  $x, y \in \mathbb{R}$ ,  $|\sigma(y) - \sigma(x)| \leq \lambda(|x - y|)$ . Then, the following claims hold true.*

(a) *If  $(X_t)_{t \geq 0}$  admits a unique invariant distribution  $\nu$ , the distribution  $\nu_\Delta = \nu \circ (\xi \mapsto (\xi, \xi))^{-1}$  is the unique invariant measure of the duplicated diffusion  $(X_t^{x_1}, X_t^{x_2})_{t \geq 0}$ .*

(b) *(Has'minskii) Assume that the scale function  $p$  is well-defined as a real function on the real line and that,*

$$\lim_{x \rightarrow \pm\infty} p(x) = \pm\infty \quad \text{and} \quad M \text{ is finite.}$$

*Then, a.s., for every  $x_1, x_2 \in \mathbb{R}$ ,  $p(X_t^{x_1}) - p(X_t^{x_2})$  tends to 0 when  $t$  goes to infinity,*

**REMARK 2.1.**  $\triangleright$  The general assumptions on  $b$  and  $\sigma$  are obviously fulfilled whenever these functions are locally Lipschitz with linear growth.

$\triangleright$  The proofs of both claims are based on (typically one-dimensional) comparison arguments. This also explains the assumption on  $\sigma$  which is a classical sufficient assumption to ensure comparison of solutions, namely, if  $x_1 \leq x_2$ , then  $X_t^{x_1} \leq X_t^{x_2}$  for every  $t \geq 0$  a.s. (see [IW77]).

$\triangleright$  Note that the assumptions in (b) imply the uniqueness of  $\nu$ : when the speed measure  $M$  is finite and  $\sigma$  never vanishes, the distribution  $\nu(d\xi) = M(d\xi)/M(\mathbb{R})$  is the unique invariant measure of the diffusion. Thus, by (a), we also have the uniqueness of the invariant distribution for the duplicated diffusion. However, in this case, it can be viewed as a direct consequence of the asymptotic confluence of  $p(X_t^{x_1})$  and  $p(X_t^{x_2})$  as  $t \rightarrow +\infty$ . Actually, if for all  $x_1, x_2 \in \mathbb{R}^d$ ,  $p(X_t^{x_1}) - p(X_t^{x_2}) \xrightarrow{t \rightarrow +\infty} 0$  a.s, we deduce that for any  $K > 0$

$$\int_{\mathbb{R}} \left( |p(x_1) - p(x_2)| \wedge K \right) \mu(dx_1, dx_2) \leq \limsup_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \mathbb{E}_\mu \left( |p(X_s^{x_1}) - p(X_s^{x_2})| \wedge K \right) ds = 0.$$

As a consequence,  $p(x_1) = p(x_2)$   $\mu(dx_1, dx_2)$ -a.s. Since  $p$  is an increasing function, it follows that  $\mu(\{(x, x), x \in \mathbb{R}\}) = 1$  and thus that  $\mu = \nu_\Delta$ .

▷ As mentioned before, (b) slightly extends a result by Has'minkii obtained in [Has80] with different methods and under the additional assumption that  $\sigma$  never vanishes (whereas we only need the scale function  $p$  to be finite which allows e.g. for the existence of integrable singularities of  $\frac{b}{\sigma^2}$ ). Note however that the case of an infinite speed measure  $M$  (which corresponds to null recurrent diffusions) is also investigated in [Has80], requiring extra non-periodicity assumptions on  $\sigma$ .

*Proof.* (a) Throughout the proof we denote by  $(X_t^{x_1}, X_t^{x_2})$  the duplicated diffusion at time  $t \geq 0$  and by  $(Q_t((x_1, x_2), dy_1, dy_2))_{t \geq 0}$  its Feller Markov semi-group. The set  $\mathcal{I}_{DSDS}$  of invariant distributions of  $(Q_t)_{t \geq 0}$  is clearly nonempty, convex and weakly closed. Since any such distribution  $\mu$  has  $\nu$  as marginals (in the sense  $\mu(dx_1 \times \mathbb{R}) = \mu(\mathbb{R} \times dx_2) = \nu$ ), the set  $\mathcal{I}_{DSDS}$  is tight and consequently weakly compact in the the topological vector space of signed measures on  $(\mathbb{R}^2, \mathcal{B}or(\mathbb{R}^2))$  equipped with the weak topology. As a consequence of the Krein-Millman Theorem,  $\mathcal{I}_{DSDS}$  admits extremal distributions and is the convex hull of these extremal distributions.

Let  $\mu$  be such an extremal distribution and consider the following three subsets of  $\mathbb{R}^2$ :

$$A^+ = \{(x_1, x_2), x_2 > x_1\}, A^- = \{(x_1, x_2), x_1 > x_2\} \text{ and } A_0 = \{(x, x), x \in \mathbb{R}\} = \Delta_{\mathbb{R}^2}.$$

We first want to show that if  $\mu(A^+) > 0$  then the conditional distribution  $\mu^{A^+}$  defined by  $\mu^{A^+} = \frac{\mu(\cdot \cap A^+)}{\mu(A^+)}$  is also an invariant distribution for  $(Q_t)_{t \geq 0}$ . Under the above assumptions on  $b$  and  $\sigma$ , one derives from classical comparison theorems and strong pathwise uniqueness arguments for the solutions of (SDE) (see e.g. [IW77]) that

$$\forall (x_1, x_2) \in {}^c A^+ = \mathbb{R}^2 \setminus A^+, \quad Q_t((x_1, x_2), {}^c A^+) = 1.$$

We deduce that for every  $(x_1, x_2) \in \mathbb{R}^2$  and  $t \geq 0$ ,

$$Q_t((x_1, x_2), A^+) = \mathbb{P}((X_t^{x_1}, X_t^{x_2}) \in A^+) = \mathbf{1}_{A^+}(x_1, x_2) \mathbb{P}(\tau_{x_1, x_2} > t)$$

where  $\tau_{x_1, x_2} = \inf\{t \geq 0, X_t^{x_2} \leq X_t^{x_1}\}$ . The second equality follows from the pathwise uniqueness since no bifurcation can occur. Now, let  $\mu \in \mathcal{I}_{DSDS}$ . Integrating the above equality and letting  $t$  go to infinity implies

$$\mu(A^+) = \int_{A^+} \mu(dx_1, dx_2) \mathbb{P}(\tau_{x_1, x_2} = +\infty).$$

If  $\mu(A^+) > 0$ , then  $\mu(dx_1, x_2)$ -a.s.  $\mathbb{P}(\tau_{x_1, x_2} = +\infty) = 1$  on  $A^+$  i.e.  $X_t^{x_2} > X_t^{x_1}$  for every  $t \geq 0$  a.s.. As a consequence,  $\mu(dx_1, dx_2)$ -a.s., for every  $B \in \mathcal{B}(\mathbb{R}^{2d})$ ,

$$\mathbf{1}_{(x_1, x_2) \in A^+} Q_t((x_1, x_2), B) = \mathbf{1}_{(x_1, x_2) \in A^+} Q_t((x_1, x_2), B \cap A^+) = Q_t((x_1, x_2), B \cap A^+)$$

where we used again that  $Q_t((x_1, x_2), A^+) = 0$  if  $x_2 \leq x_1$ . Then, since  $\mu$  is invariant, we deduce from an integration of the above equality that

$$\mu(B \cap A^+) = \int_{\mathbb{R}^2} Q_t((x_1, x_2), B) \mathbf{1}_{(x_1, x_2) \in A^+} \mu(dx_1, dx_2).$$

It follows that if  $\mu(A^+) > 0$ ,  $\mu^{A^+}$  is invariant.

If  $\mu(A^+) < 1$ , one shows likewise that  $\mu^{cA^+}$  an invariant distribution for  $(Q_t)_{t \geq 0}$  as well. Then, if  $\mu(A^+) \in (0, 1)$ , then  $\mu$  is a convex combination of elements of  $\mathcal{I}_{DSDS}$

$$\mu = \mu(A^+) \mu^{A^+} + \mu({}^c A^+) \mu^{cA^+}$$

so that  $\mu$  cannot be extremal. Finally  $\mu(A^+) = 0$  or 1.

Assume  $\mu(A^+) = 1$  so that  $\mu = \mu(\cdot \cap A^+)$ . This implies that  $X_0^1 > X_0^2$   $\mathbb{P}_\mu$ -a.s.. But  $\mu$  being invariant, both its marginals are  $\nu$  i.e.  $X_0^1$  and  $X_0^2$  are  $\nu$ -distributed. This yields a contradiction. Indeed, let  $\varphi$  be a bounded increasing positive function. For instance, set  $\varphi(u) := 1 + \frac{u}{\sqrt{u^2+1}}$ ,  $u \in \mathbb{R}$ . Then,  $\mathbb{E}[\varphi(X_0^1) - \varphi(X_0^2)] > 0$  since  $X_0^1 > X_0^2$   $\mathbb{P}_\mu$ -a.s. but we also have  $\mathbb{E}[\varphi(X_0^1) - \varphi(X_0^2)] = 0$  since  $X_0^1$  and  $X_0^2$  have the same distribution. This contradiction implies that  $\mu(A^+) = 0$ .

One shows likewise that  $\mu(A^-) = 0$  if  $\mu$  is an extremal measure. Finally any extremal distribution of  $\mathcal{I}_{DSDS}$  is supported by  $A_0 = \Delta_{\mathbb{R}^2}$ . Given the fact that the marginals of  $\mu$  are  $\nu$  this implies that  $\mu = \nu_\Delta = \nu \circ (x \mapsto (x, x))^{-1}$  which in turn implies that  $\mathcal{I}_{DSDS} = \{\nu_\Delta\}$ .

(b) Let  $x_1, x_2 \in \mathbb{R}$ . If  $x_1 > x_2$  then  $X_t^{x_1} \geq X_t^{x_2}$ , still by a comparison argument, and  $p(X_t^{x_1}) \geq p(X_t^{x_2})$  since  $p$  is increasing. Consequently  $M_t^{x_1, x_2} = p(X_t^{x_1}) - p(X_t^{x_2})$ ,  $t \geq 0$ , is a non-negative continuous local martingale, hence  $\mathbb{P}$ -a.s. converging toward a finite random limit  $\ell_\infty^{x_1, x_2} \geq 0$ . One proceeds likewise when  $x_1 < x_2$  (with  $\ell_\infty^{x_1, x_2} \leq 0$ ). When  $x_1 = x_2$ ,  $M_t = \ell_\infty^{x_1, x_2} \equiv 0$ . The aim is now to show that  $\ell_\infty^{x_1, x_2} = 0$  a.s. To this end, we introduce

$$\mu_t(dy_1, dy_2) := \frac{1}{t} \int_0^t Q_s((x_1, x_2), dy_1, dy_2) ds, \quad (x_1, x_2) \in \mathbb{R}^{2d}$$

and we want to check that for every  $(x_1, x_2) \in \mathbb{R}^{2d}$ ,  $(\mu_t(dy_1, dy_2))_{t \geq 1}$  converges weakly to  $\nu_\Delta$ . Owing to the uniqueness of  $\nu_\Delta$  established in (a) and to the fact that any weak limiting distribution of  $(\mu_t(dy_1, dy_2))_{t \geq 1}$  is always invariant (by construction), it is enough to prove that  $(\mu_t(dy_1, dy_2))_{t \geq 1}$  is tight. Since the tightness of a sequence of probability measures defined on a product space is clearly equivalent to that of its first and second marginals, it is here enough to prove the tightness of  $(t^{-1} \int_0^t P_s(x_0, dy) ds)_{t \geq 1}$  for any  $x_0 \in \mathbb{R}$ .

Let  $x_0 \in \mathbb{R}$ . Owing to the comparison theorems, we have for all  $t \geq 0$  and  $M \in \mathbb{R}$ ,  $P_t(x_0, [M, +\infty)) \leq P_t(x, [M, +\infty))$  if  $x \geq x_0$  and  $P_t(x_0, (-\infty, M]) \leq P_t(x, (-\infty, M])$  if  $x_0 \geq x$ . Since  $\nu$  is invariant and equivalent to the Lebesgue measure, we deduce that

$$P_t(x_0, [M, +\infty)) \leq \frac{\nu([M, +\infty))}{\nu([x_0, +\infty))} \quad \text{and} \quad P_t(x_0, (-\infty, M]) \leq \frac{\nu((-\infty, M])}{\nu((-\infty, x_0])}$$

The tightness of  $(P_t(x_0, dy))_{t \geq 1}$  follows (from that of  $\nu$ ) and we derive from what precedes that

$$\forall (x_1, x_2) \in \mathbb{R}^{2d}, \quad \frac{1}{t} \int_0^t Q_s((x_1, x_2), dy_1, dy_2) ds \xrightarrow{(\mathbb{R}^d)} \nu_\Delta(dy_1, dy_2).$$

Now, note that for every  $L \in \mathbb{N}$ , the function  $g_L : (y_1, y_2) \mapsto |p(y_1) - p(y_2)| \wedge L$  is continuous and bounded. Hence by Césaro's Theorem, we have that

$$\frac{1}{t} \int_0^t Q_s(g_L)(x_1, x_2) ds = \frac{1}{t} \int_0^t \mathbb{E} g_L(X_s^{x_1}, X_s^{x_2}) ds \longrightarrow \mathbb{E}(|\ell_\infty^{x_1, x_2}| \wedge L)$$

whereas, by the above weak convergence of  $(\mu_t(dy_1, dy_2))_{t \geq 1}$ , we get

$$\frac{1}{t} \int_0^t Q_s(g_L)(x_1, x_2) ds \longrightarrow \int_{\mathbb{R}^d} g_L(y_1, y_2) \nu_\Delta(dy_1, dy_2) = 0 \quad \text{as} \quad t \rightarrow +\infty$$

since  $g_L$  is identically 0 on  $\Delta_{\mathbb{R}^{2d}}$ . It follows, by letting  $L$  go to infinity, that

$$\mathbb{E} |\ell_\infty^{x_1, x_2}| = 0$$



This implies  $\ell_\infty^{x_1, x_2} = 0$   $\mathbb{P}$ -a.s. which in turn implies that

$$\mathbb{P}\text{-a.s. } p(X_t^{x_1}) - p(X_t^{x_2}) \longrightarrow 0 \quad \text{as } t \rightarrow +\infty.$$

Finally, it remains to prove that we can exchange the quantifiers, *i.e.* that  $\mathbb{P}$ -a.s.,  $p(X_t^{x_1}) - p(X_t^{x_2}) \longrightarrow 0$  for every  $x_1, x_2$ . Assume that  $x_1 \geq x_2$ . Again by the comparison theorem and the fact that  $p$  increases, we have  $0 \leq p(X_t^{x_1}) - p(X_t^{x_2}) \leq p(X_t^{\lfloor x_1 \rfloor + 1}) - p(X_t^{\lfloor x_2 \rfloor})$ . This means that we can come down to a countable set of starting points.  $\square$

In the continuity of the second part of Theorem 2.1(b), it is natural to wonder whether a one-dimensional diffusion is asymptotically confluent, *i.e.* when for all  $x_1, x_2 \in \mathbb{R}$ ,  $X_t^{x_1} - X_t^{x_2}$  tends to 0 *a.s.* as  $t \rightarrow +\infty$ . In the following corollary, we show that such property holds in a quite general setting.

**COROLLARY 2.1.** (a) Assume the hypothesis of Theorem 2.1(b) hold. If furthermore,

$$\sigma \text{ never vanishes} \quad \text{and} \quad \limsup_{|x| \rightarrow +\infty} \int_0^x \frac{b}{\sigma^2}(\xi) d\xi < +\infty$$

then,  $\mathbb{P}$ -a.s., for every  $x_1, x_2 \in \mathbb{R}$ ,

$$X_t^{x_1} - X_t^{x_2} \longrightarrow 0 \quad \text{as } t \rightarrow +\infty.$$

(b) The above condition is in particular satisfied if there exists  $M > 0$  such that for all

$$|x| > M \implies \text{sign}(x)b(x) \leq 0.$$

**REMARK 2.2.** Let  $U$  be a positive a twice differentiable function such that  $\lim_{|x| \rightarrow +\infty} U(x) = +\infty$  and consider the one-dimensional Kolmogorov equation

$$dX_t = -U'(X_t)dt + \sigma dW_t$$

where  $\sigma > 0$ . Then,

$$\liminf_{|x| \rightarrow +\infty} xU'(x) > \frac{\sigma^2}{2} \implies X_t^x - X_t^y \xrightarrow{t \rightarrow +\infty} 0 \text{ a.s.}$$

Note that in particular, this result holds true even if  $U$  has several local minimas.

*Proof.* (a) Under the assumptions of the theorem,  $p$  is continuously differentiable on  $\mathbb{R}$  and

$$p'(x) = e^{-\int_{x_0}^x \frac{2b}{\sigma^2}(u) du}, \quad x \in \mathbb{R}.$$

Then it is clear that  $p'_{\min} = \inf_{x \in \mathbb{R}} p'(x) > 0$  iff  $\limsup_{|x| \rightarrow +\infty} \int_{x_0}^x \frac{2b}{\sigma^2}(\xi) d\xi < +\infty$ . By the fundamental theorem of calculus, we know that,

$$|X_t^{x_1} - X_t^{x_2}| \leq \frac{1}{p'_{\min}} |p(X_t^{x_1}) - p(X_t^{x_2})|,$$

and the result follows from Theorem 2.1(b).

(b) Since  $\sigma$  never vanishes,  $p''$  is well-defined and for every  $x \in \mathbb{R}$ ,  $p''(x) = -\frac{2b(x)p'(x)}{\sigma^2(x)}$ . Using that  $p'$  is positive, we deduce from the assumptions that

$$\exists M > 0 \quad \text{such that} \quad \begin{cases} p''(x) \geq 0, & x \geq M \\ p''(x) \leq 0, & x \leq -M. \end{cases}$$

Now,  $p'$  being continuous, it follows that  $p'$  attains a positive minimum  $p'_{\min} > 0$ .  $\square$

### 3 The multidimensional case

In this section, we begin by an example of a multidimensional Brownian diffusion  $(X^{x_1}, X^{x_2})$  for which  $\nu_\Delta$  (image of  $\nu$  on the diagonal) is not the only one invariant distribution. Thus, Theorem 2.1 is specific to the case  $d = 1$  and we can not hope to get a similar result for the general case  $d \geq 2$ .

#### 3.1 Counterexample in 2-dimension

Roughly speaking, saying that  $\nu_\Delta$  is the only one invariant distribution means in a sense that  $X_t^x - X_t^y$  has a tendency to converge towards 0 when  $t \rightarrow +\infty$ . Thus, the idea in the counterexample below is to build a “turning” two-dimensional ergodic process where the angular difference between the two coordinates does not depend on  $t$ . Such a construction leads to a model where the distance between the two coordinates can not tend to 0 (Note that some proofs are deferred to Appendix B).

We consider the 2-dimensional *SDE* with Lipschitz continuous coefficients defined by,  $\forall x \in \mathbb{R}^2$

$$\begin{aligned} b(x) &= \left( x \mathbf{1}_{\{0 \leq |x| \leq 1\}} - \frac{x}{|x|} \mathbf{1}_{\{|x| \geq 1\}} \right) (1 - |x|) \\ \sigma(x) &= \vartheta \text{Diag}(b(x)) + \begin{bmatrix} 0 & -cx^2 \\ cx^1 & 0 \end{bmatrix}. \end{aligned}$$

where  $\vartheta, c \in (0, +\infty)$  are fixed parameters.

Switching to polar coordinates  $X_t = (r_t \cos \varphi_t, r_t \sin \varphi_t)$ ,  $t \in \mathbb{R}_+$ , we obtain that this *SDE* also reads

$$dr_t = \min(r_t, 1)(1 - r_t)(dt + \vartheta dW_t^1), \quad r_0 \in \mathbb{R}_+ \quad (3.3)$$

$$d\varphi_t = c dW_t^2, \quad \varphi_0 \in [0, 2\pi), \quad (3.4)$$

where  $x_0 = r_0(\cos \varphi_0, \sin \varphi_0)$  and  $W = (W^1, W^2)$  is a standard 2-dimensional Brownian motion.

Standard considerations about Feller classification (see Appendix B for details) show that, if  $x_0 \neq 0$  (*i.e.*  $r_0 > 0$ ) and  $\vartheta \in (0, \sqrt{2})$  then

$$r_t \longrightarrow 1 \quad \text{as } t \rightarrow +\infty, \quad (3.5)$$

while it is classical background that

$$\mathbb{P}\text{-a.s. } \forall \varphi_0 \in \mathbb{R}_+, \quad \frac{1}{t} \int_0^t \delta_{e^{i(\varphi_0 + cW_s^2)}} ds \Longrightarrow \lambda_{\mathcal{S}_1} \quad \text{as } t \rightarrow +\infty$$

where  $\mathcal{S}_1$  denotes the unit circle of  $\mathbb{R}^2$ . Combining these two results straightforwardly yields

$$\forall x \in \mathbb{R}^2 \setminus \{(0, 0)\}, \quad \mathbb{P}\text{-a.s. } \frac{1}{t} \int_0^t \delta_{X_s^x} ds \xrightarrow{(\mathbb{R}^2)} \lambda_{\mathcal{S}_1} \quad \text{as } t \rightarrow +\infty.$$

On the other hand, given the form of  $\varphi_t$ , it is clear that if  $x = r_0 e^{i\varphi_0}$  and  $x' = r'_0 e^{i\varphi'_0}$ ,  $r_0, r'_0 \neq 0$ ,  $\varphi_0, \varphi'_0 \in [0, 2\pi)$ , then

$$\lim_{t \rightarrow +\infty} |X_t^x - X_t^{x'}| = |e^{i(\varphi_0 - \varphi'_0)} - 1|$$

which in turn implies that

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t |X_s^x - X_s^{x'}| ds = |e^{i(\varphi_0 - \varphi'_0)} - 1| \quad \mathbb{P}\text{-a.s.}$$

This limit being different from 0 as soon as  $\varphi_0 \neq \varphi'_0$ , one derives, as a consequence, that  $\nu_\Delta$  cannot be the only invariant distribution. In fact, a more precise statement can be proved.

**PROPOSITION 3.1.** (a) *A distribution  $\mu$  is invariant for the semi-group  $(Q_t)_{t \geq 0}$  of the duplicated diffusion if and only if  $\mu$  has the following form:*

$$\mu = \mathcal{L}(e^{i\Theta}, e^{i(\Theta+V)}), \quad (3.6)$$

where  $\Theta$  is uniformly distributed over  $[0, 2\pi]$  and  $V$  is a  $[0, 2\pi)$ -valued random variable independent of  $\Theta$

(b) *When  $V = 0$  a.s., we retrieve  $\nu_\Delta$  whereas, when  $V$  also has uniform distribution on  $[0, 2\pi]$ , we obtain  $\nu \otimes \nu$ . Finally,  $\mu$  is extremal in the convex set of  $(Q_t)_{t \geq 0}$  invariant distributions if and only if there exists  $\theta_0 \in [0, 2\pi)$  such that  $V = \theta_0$  a.s.*

The proof is postponed to Appendix B. However, note that the claim about extremal invariant distributions follows from the fact that for every  $\theta \in [0, 2\pi)$ ,  $(Q_t)_{t \geq 0}$  leaves the set  $\Gamma_\theta := \{(e^{i\varphi}, e^{i\varphi'}) \in \mathcal{S}_1 \times \mathcal{S}_1, \varphi' - \varphi \equiv \theta \text{ mod. } 2\pi\}$  stable.

**REMARK 3.3.** In the above counterexample, the invariant measure of  $(r_t)_{t \geq 0}$  is the Dirac measure  $\delta_1$ . In fact, setting again  $x = \varepsilon_0 e^{i\varphi_0}$  and  $x' = r'_0 e^{i\varphi'_0}$  and using that  $X_t^x - X_t^{x'} = r_t^x (e^{i(\varphi_0 + W_t^2)} - e^{i(\varphi'_0 + W_t^2)}) + (r_t^x - r_t^{x'}) e^{i(\varphi'_0 + W_t^2)}$ , an easy adaptation of the above proof shows that it can be generalized to any ergodic non-negative process  $(r_t)_{t \geq 0}$  solution to an autonomous SDE and satisfying the following properties:

- Its unique invariant distribution  $\pi$  satisfies  $\pi(\mathbb{R}_+^*) = 1$ .
- For every  $x, y \in (0, +\infty)$ ,  $r_t^x - r_t^y \rightarrow 0$  a.s. as  $t \rightarrow +\infty$ .

For instance, let  $(X_t^x)_{t \geq 0}$  be an Ornstein-Uhlenbeck process satisfying the SDE  $dX_t = -X_t dt + \sigma dW_t$ ,  $X_0 = x$ . Set  $r_t^x = (X_t^x)^2$  (this is a special case of the Cox-Ingersoll-Ross process). The process  $(r_t^x)$  clearly satisfies the first two properties. Furthermore,  $(X_t^x)_{t \geq 0}$  satisfies a.s. for every  $x, y \in \mathbb{R}$  and every  $t \geq 0$ ,  $|X_t^x - X_t^y| = |x - y|e^{-t}$ . Then, since for every  $x \in \mathbb{R}$ ,

$$\frac{X_t^x}{t} = -\frac{1}{t} \int_0^t X_s^x ds + \sigma \frac{W_t}{t} \rightarrow 0 \quad \text{a.s.} \quad \text{as} \quad t \rightarrow +\infty,$$

it follows that  $(r_t^x)_{t \geq 0}$  also satisfies for all positive  $x, y$ ,  $r_t^x - r_t^y \rightarrow 0$  a.s. as  $t \rightarrow +\infty$  (Many other examples can be built using Corollary 2.1).

Finally, note that if  $\mu = \mathcal{L}(Re^{i\Theta}, Re^{i(\Theta+V)})$  where  $R$ ,  $\Theta$  and  $V$  are independent random variables such that the distributions of  $R$  and  $\Theta$  are respectively  $\pi$  and the uniform distribution on  $[0, 2\pi]$  and  $V$  takes values in  $[0, 2\pi)$ , then  $\mu$  is an invariant distribution of the associated duplication system.

**REMARK 3.4.** In relation to this counterexample we can mention a general result on the Brownian flows of Harris (see [Har81],[Kun90] Theorem 4.3.2). The theorem gives conditions on  $b$  and  $\sigma$  under which  $\nu$  is an invariant measure of the one point motion  $(X_t^x)_{t \geq 0}$  and  $\nu \otimes \nu$  is an invariant measure of the two point motion  $(X_t^{x_1}, X_t^{x_2})_{t \geq 0}$ .

Harris's Theorem also allows to elucidate the case of linear systems. More precisely, let  $(X_t^x)_{t \geq 0}$  be the solution to the linear system

$$dX_t^x = AX_t^x dt + \sum_{k=1}^q \Sigma_k X_t^x dW_t^k, \quad X_0^x = x \in \mathbb{R}^d,$$

where  $W = (W^1, \dots, W^q)$  is a  $q$ -dimensional standard Brownian motion and  $A, \Sigma_\ell \in \mathcal{M}(d, d)$ ,  $\ell = 1, \dots, q$  are constant matrices. The Lyapunov exponent  $\lambda$  of the system is given by  $\text{Tr}A - \frac{1}{2} \sum_{k=1}^q \text{Tr}(\Sigma_k \Sigma_k^*)$ .

– If  $\lambda < 0$ ,  $X_t^x \rightarrow 0$  as  $t \rightarrow \infty$  and if  $\lambda > 0$ ,  $|X_t^x| \rightarrow +\infty$  as  $t \rightarrow \infty$  so that no steady regime can exist.

– If  $\lambda = 0$ , the Lebesgue measure  $\lambda_d$  on  $\mathbb{R}^d$  is invariant for  $(X_t^x)_{t \geq 0}$  and both the Lebesgue measure  $\lambda_{2d} = \lambda_d \otimes \lambda_d$  on  $\mathbb{R}^{2d}$  and  $\lambda_\Delta$  are invariant for the duplicated system  $(X_t^{x_1}, X_t^{x_2})_{t \geq 0}$ .

### 3.2 Uniqueness of the invariant measure: $\theta$ -confluence

In the sequel of this section, we build some criterions for the uniqueness of the invariant distribution of the duplicated system in the multidimensional case. The main idea of our criterions discussed below is to analyze the coupled diffusion process  $(X^{x_1}, X^{x_2})$  through the squared distance process  $r_t = |X^{x_1} - X^{x_2}|^2$ . This idea is similar to that of Has'minskii's test for explosion of diffusions in  $\mathbb{R}^d$  and to the one proposed in Chen and Li's work devoted to the coupling of diffusions (see [CL89]).

**THEOREM 3.2.** *Assume the SDE (1.1) is weakly asymptotically  $\theta$ -confluent in the following sense: there exists a continuous function  $\theta : (0, +\infty) \rightarrow \mathbb{R}$  such that*

$$\left\{ \begin{array}{l} (i) \quad \int_0^1 e^{2 \int_v^1 \frac{\theta(w)}{w} dw} dv < +\infty. \\ (ii) \quad \forall x, y \in \mathbb{R}^d, x \neq y, \\ \quad (b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2 < \theta(|x - y|^2) \left| (\sigma^*(x) - \sigma^*(y)) \frac{x - y}{|x - y|} \right|^2, \\ \text{or} \\ (ii') \quad \forall x, y \in \mathbb{R}^d, x \neq y, \\ \quad \begin{cases} (\sigma^*(x) - \sigma^*(y))(x - y) \neq 0 \\ (b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2 \leq \theta(|x - y|^2) \left| (\sigma^*(x) - \sigma^*(y)) \frac{x - y}{|x - y|} \right|^2. \end{cases} \end{array} \right. \quad (3.7)$$

(a) Uniqueness of invariant distribution(s): *Assume that the set  $\mathcal{I}_{SDE}$  of invariant distributions of SDE is (nonempty, convex and) weakly compact. Then, if  $\mathcal{I}_{DSDS}$  denotes the set of invariant distributions of the duplicated system (DSDS), one has*

$$\mathcal{I}_{SDE} = \{\nu\} \quad \text{and} \quad \mathcal{I}_{DSDS} = \left\{ \nu \circ (x \mapsto (x, x))^{-1} \right\}.$$

(b) Pathwise confluence: *If furthermore, for every  $x \in \mathbb{R}^d$ ,  $\left( \frac{1}{t} \int_0^t P_s(x, dy) ds \right)_{t \geq 0}$  is tight (so that the set  $\mathcal{I}_{SDE}$  of invariant distributions of  $(P_t)_{t \geq 0}$  is not empty and weakly compact), we have a.s. pathwise asymptotic confluence:*

$$\forall x_1, x_2 \in \mathbb{R}^d, \quad X_t^{x_1} - X_t^{x_2} \longrightarrow 0 \text{ P-a.s.} \quad (3.8)$$

**REMARK 3.5.** If  $b$  and  $\sigma$  are Lipschitz continuous, Kunita's Stochastic flow theorem (see [Kun90], Section 4.5) ensures in particular that, if  $x_1 \neq x_2$ , the solutions  $X_t^{x_1}$  and  $X_t^{x_2}$  *a.s.* never get stuck. Taking advantage of this remark slightly shortens the proof below.

**REMARK 3.6.** Criteria for the tightness of  $\left(\frac{1}{t} \int_0^t P_s(x, dy) ds\right)_{t \geq 0}$  for every  $x \in \mathbb{R}^d$  usually rely on the mean-reversion property of the solutions of  $(SDE)$  usually established under various assumptions involving the existence of a so-called *Lyapunov function*  $V$  going to infinity at infinity and such that  $\mathcal{A}V$  is upper-bounded and  $\limsup_{|x| \rightarrow +\infty} \mathcal{A}V(x) < 0$  where  $\mathcal{A}$  denotes the infinitesimal generator of  $X^x$  (so-called *Has'minskii's criterion*). Keep in mind that

$$\mathcal{A}V(x) = (b|\nabla V)(x) + \frac{1}{2}\text{Tr}\left(\sigma\sigma^*(x)D^2V(x)\right)$$

where  $\text{Tr}(A)$  stands for the trace of the matrix  $A$ .

On the other hand, a classical criterion for pathwise asymptotic confluence (at exponential rate, see *e.g.* [BB92], [Lem05]) is

$$\forall x, y \in \mathbb{R}^d, \quad (b(x) - b(y)|x - y) + \frac{1}{2}\|\sigma(x) - \sigma(y)\|^2 < -c|x - y|^2 \quad (c > 0),$$

and, as a straightforward consequence, uniqueness of the invariant distribution  $\nu$  of  $(SDE)$  (and of  $(DSDS)$  as well). Moreover, putting  $y = 0$  in the above inequality straightforwardly yields real coefficients  $\alpha > 0$ ,  $\beta \geq 0$  such that  $\mathcal{A}V \leq \beta - \alpha V$  with  $V(x) = |x|^2$ . Hence Has'minskii criterion is fulfilled, so it is also an existence criterion for the invariant distribution which  $\theta$ -confluence is not.

But in fact, the  $\theta$ -confluence criterion turns out to be a much weaker condition to ensure the uniqueness of the invariant distribution  $\nu$  of  $(SDE)$  (and of  $\nu_\Delta$  for  $(DSDS)$ ). This is emphasized by several operating criteria derived from this  $\theta$ -confluence assumptions. Finally, combined with the tightness of the occupation measure of the semi-group, it becomes a criterion for *a.s.* pathwise asymptotic confluence.

*Proof.* STEP 1: Exactly like in the beginning of the proof of Theorem 2.1(a), one checks that the set  $\mathcal{I}_{DSDS}$  of invariant distributions of  $(Q_t)_{t \geq 0}$  is a nonempty, convex and weakly compact subset of  $\mathcal{P}(\mathbb{R}^{2d})$ . As a consequence of the Krein-Millman theorem,  $\mathcal{I}_{DSDS}$  admits extremal distributions.

On the other hand, under the above Lipschitz assumptions on  $b$  and  $\sigma$ , it follows from strong uniqueness theorem for SDE's that the semi-group  $(Q_t)_{t \geq 0}$  leaves stable the diagonal  $\Delta_{\mathbb{R}^{2d}}$ .

Let  $x_1, x_2 \in \mathbb{R}^d$ ,  $x_1 \neq x_2$ . We define the stopping time

$$\tau_{x_1, x_2} := \inf \{t \geq 0 \mid X_t^{x_1} = X_t^{x_2}\}.$$

Still by a strong uniqueness argument it is clear that  $\{\tau_{x_1, x_2} > t\} = \{X_t^{x_1} \neq X_t^{x_2}\}$  so that

$$Q_t((x_1, x_2), {}^c\Delta_{\mathbb{R}^{2d}}) = \mathbf{1}_{{}^c\Delta_{\mathbb{R}^{2d}}}(x_1, x_2)\mathbb{P}(\tau_{x_1, x_2} > t)$$

and  $Q_t((x_1, x_1), {}^c\Delta_{\mathbb{R}^{2d}}) = 0$ .

Let  $\mu \in \mathcal{I}_{DSDS}$  be an extremal invariant measure. We have, for every  $t \geq 0$ ,

$$\mu({}^c\Delta_{\mathbb{R}^{2d}}) = \int_{{}^c\Delta_{\mathbb{R}^{2d}}} \mu(dx_1, dx_2)\mathbb{P}(\tau_{x_1, x_2} > t).$$

Letting  $t$  go to  $+\infty$  yields

$$\mu({}^c\Delta_{\mathbb{R}^{2d}}) = \int_{{}^c\Delta_{\mathbb{R}^{2d}}} \mu(dx_1, dx_2) \mathbb{P}(\tau_{x_1, x_2} = +\infty)$$

so that, on  ${}^c\Delta_{\mathbb{R}^{2d}}$ ,  $\mu(dx_1, dx_2)$ -a.s.,  $\mathbb{P}(\tau_{x_1, x_2} = +\infty) = 1$  or equivalently the process  $(X^{x_1}, X^{x_2})$  lives in  ${}^c\Delta_{\mathbb{R}^{2d}}$ . Consequently, if  $\mu({}^c\Delta_{\mathbb{R}^{2d}}) \in (0, 1)$ , both conditional measures  $\mu^{{}^c\Delta_{\mathbb{R}^{2d}}}$  and  $\mu^{\Delta_{\mathbb{R}^{2d}}}$  are invariant distributions for  $(SDSD)$  as well. If so,

$$\mu = \mu({}^c\Delta_{\mathbb{R}^{2d}})\mu^{{}^c\Delta_{\mathbb{R}^{2d}}} + \mu(\Delta_{\mathbb{R}^{2d}})\mu^{\Delta_{\mathbb{R}^{2d}}}$$

cannot be extremal. Consequently  $\mu(\Delta_{\mathbb{R}^{2d}}) = 0$  or  $1$ .

STEP 2: Let  $\mu$  be an extremal distribution in  $\mathcal{I}_{SDSD}$  such that  $\mu({}^c\Delta_{\mathbb{R}^{2d}}) = 1$ . Let  $v_0 \in (0, +\infty)$  any (fixed) real number. Let  $f : [0, +\infty) \rightarrow \mathbb{R}_+$  be the function null at  $0$  and twice continuously differentiable on  $(0, +\infty)$  defined by

$$f(\xi) = \int_0^\xi e^{2 \int_v^{v_0} \frac{\theta(w)}{w} dw} dv.$$

The function  $f$  satisfies the  $ODE \equiv 2\theta(\xi)f'(\xi) + \xi f''(\xi) = 0$ ,  $\xi \in (0, +\infty)$  and  $f' > 0$  on  $(0, +\infty)$ . We know from Step 1 that  $\mu(dx_1, dx_2)$ -a.s.,  $(X_t^{x_1} \neq X_t^{x_2}$  for every  $t \geq 0$ ) a.s. so that we can apply Itô's formula to the duplicated system  $(X^{x_1}, X^{x_2})$  with the function  $\varphi(y_1, y_2) := f(|y_1 - y_2|^2)$ :

$$\begin{aligned} \varphi(X_t^{x_1}, X_t^{x_2}) &= \varphi(x, y) + \int_0^t \mathcal{A}^{(2)}\varphi(X_s^{x_1}, X_s^{x_2}) ds \\ &\quad + \underbrace{\int_0^t f'(|X_s^{x_1} - X_s^{x_2}|^2) ((\sigma^*(X_s^{x_1}) - \sigma^*(X_s^{x_2}))(X_s^{x_1} - X_s^{x_2}) |dW_s)}_{=: M_t \text{ local martingale}} \end{aligned}$$

where

$$\begin{aligned} \mathcal{A}^{(2)}\varphi(x_1, x_2) &= 2 \left( (b(x_1) - b(x_2)) |x_1 - x_2| + \frac{1}{2} \|\sigma(x_1) - \sigma(x_2)\|^2 \right) f'(|x_1 - x_2|^2) \\ &\quad + f''(|x_1 - x_2|^2) |(\sigma^*(x_1) - \sigma^*(x_2))(x_1 - x_2)|^2. \end{aligned}$$

If (ii) (resp. (ii')) holds true, it follows from the definition of  $f$  that  $\mathcal{A}^{(2)}\varphi(x_1, x_2) < 0$ ,  $x_1 \neq x_2$  (resp.  $\leq 0$ ).

As a consequence  $(\varphi(X_t^{x_1}, X_t^{x_2}))_{t \geq 0}$  is a non-negative  $\mathbb{P}_\mu$ -supermartingale, hence it converges toward a non-negative integrable random variable  $L_\infty$  (there is here an abuse of notation since under  $\mathbb{P}_\mu$  the starting value of the duplicated system is  $\mu$ -distributed independently of  $W$ ). By localization of the local martingale, it classically follows that

$$\mathbb{E}_\mu \left( \int_0^{+\infty} (-\mathcal{A}^{(2)}\varphi(X_s^{x_1}, X_s^{x_2})) ds \right) \leq \mathbb{E} L_\infty < +\infty.$$

Then

$$\int_0^{+\infty} (-\mathcal{A}^{(2)}\varphi(X_s^{x_1}, X_s^{x_2})) ds < +\infty \quad \mathbb{P}_\mu\text{-a.s.}$$

which in turn implies

$$\frac{1}{t} \int_0^t (-\mathcal{A}^{(2)}\varphi(X_s^{x_1}, X_s^{x_2})) ds \longrightarrow 0 \quad \text{as } t \rightarrow +\infty.$$

The invariant distribution  $\mu$  being extremal,  $\mathbb{P}_\mu$  is ergodic so that by the pointwise ergodic theorem

$$\int_{\mathbb{R}^{2d} \setminus \Delta_{\mathbb{R}^{2d}}} (-\mathcal{A}^{(2)}\varphi(y_1, y_2))\mu(dy_1, dy_2) = 0.$$

This yields a contradiction under Assumption (ii).

If (ii') holds one proceeds as follows: the local martingale  $(M_t)_{t \geq 0}$  is also  $\mathbb{P}_\mu$ -a.s. converging in  $\mathbb{R}$  so that its bracket process  $\mathbb{P}_\mu$ -a.s. converges as well *i.e.*

$$\int_0^{+\infty} f'(|X_s^{x_1} - X_s^{x_2}|^2)|(\sigma^*(X_s^{x_1}) - \sigma^*(X_s^{x_2}))(X_s^{x_1} - X_s^{x_2})|^2 ds < +\infty \quad \mathbb{P}_\mu\text{-a.s.}$$

The same reasoning as above yields

$$\int_{\mathbb{R}^{2d} \setminus \Delta_{\mathbb{R}^{2d}}} f'(|x_1 - x_2|^2)|(\sigma^*(x_1) - \sigma^*(x_2))(x_1 - x_2)|^2 \mu(dx_1, dx_2) = 0$$

which implies, combined with the fact that  $f' > 0$  everywhere, that

$$(\sigma^*(x_1) - \sigma^*(x_2))(x_1 - x_2) = 0 \quad \mu(dx_1, dx_2)\text{-a.s.}$$

Under (ii') this yields a contradiction.

Consequently,  $\mu(\Delta_{\mathbb{R}^{2d}}) = 1$ . By Krein-Millman's Theorem  $\mathcal{I}_{SDS}$  is the weak closure of the convex hull of its extremal distributions. Consequently, the diagonal  $\Delta_{\mathbb{R}^{2d}}$  being a closed subset of  $\mathbb{R}^{2d}$ , all invariant distributions of the duplicated system are supported by this diagonal. For any such invariant distribution  $\mu$  both its marginals are invariant distributions for (SDE). If (SDE) had two distinct invariant distributions  $\nu$  et  $\nu'$ , we know from the introduction that  $\mathcal{I}_{SDS}$  would contain at least a distribution  $\mu$  for which the two marginals distributions are  $\mu(\cdot \times \mathbb{R}^d) = \nu$  and  $\mu(\mathbb{R}^d \times \cdot) = \nu'$  respectively. As a consequence, such a distribution  $\mu$  could not be supported by the diagonal  $\Delta_{\mathbb{R}^{2d}}$ . Finally,  $\mathcal{I}_{SDE}$  is reduced to a singleton  $\{\nu\}$  and  $\mathcal{I}_{SDS} = \{\nu \circ (\xi \mapsto (\xi, \xi))^{-1}\}$ .

STEP 3 (PROOF OF (3.8)): Let  $x_1, x_2 \in \mathbb{R}^d$ . Since  $(f(|X_t^{x_1} - X_t^{x_2}|^2))_{t \geq 0}$  is a  $\mathbb{P}$ -supermartingale, it *a.s.* converges toward  $L_\infty^{x_1, x_2} \in L^1(\mathbb{P})$ . The function  $f$  being increasing, it follows that  $|X_t^{x_1} - X_t^{x_2}|^2$  *a.s.* converges toward a finite random variable  $\ell_\infty^{x_1, x_2} = f^{-1}(L_\infty^{x_1, x_2})$ .

Now, using that for every  $x \in \mathbb{R}^d$ ,  $(\frac{1}{t} \int_0^t P_s(x, dy) ds)_{t \geq 0}$  is tight, we derive from the uniqueness of  $\nu_\Delta$  that

$$\frac{1}{t} \int_0^t Q_s((x_1, x_2), (dy_1, dy_2)) ds \xrightarrow{(\mathbb{R}^d)} \nu_\Delta$$

where  $\nu_\Delta = \nu \circ (\xi \mapsto (\xi, \xi))^{-1}$  and  $\xrightarrow{(\mathbb{R}^d)}$  denotes weak convergence of distributions on  $(\mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d))$ . Now for every bounded continuous function  $g : \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$\frac{1}{t} \int_0^t Q_s(g(|y_1 - y_2|^2))(x_1, x_2) ds = \frac{1}{t} \int_0^t \mathbb{E} g(|X_s^{x_1} - X_s^{x_2}|^2) ds \longrightarrow \mathbb{E} g(\ell_\infty^{x_1, x_2})$$

so that

$$\mathbb{E} g(\ell_\infty^{x_1, x_2}) = \int g(|y_1 - y_2|^2) \nu_\Delta(dy_1, dy_2) = g(0). \quad \square$$

### 3.3 Global criterion:

First note that Assumption (i) on the (continuous) function  $\theta$  is satisfied as soon as

$$\liminf_{u \rightarrow 0^+} \log(u) \left( \theta(u) - \frac{1}{2} \right) > \frac{1}{2}$$

so that there exists  $\kappa > 1$  and  $\varepsilon_0 \in (0, e^{-\frac{\kappa}{2}})$  such that

$$\forall u \in (0, \varepsilon_0], \quad \theta(u) \leq \frac{1}{2} \left( 1 + \frac{\kappa}{\log u} \right).$$

Note that this criterion is fulfilled in particular if there exists a  $\tilde{\kappa} \in [0, 1)$  and  $\varepsilon'_0 \in (0, +\infty)$  such that

$$\forall u \in (0, \varepsilon'_0), \quad \theta(u) \leq \frac{\tilde{\kappa}}{2}.$$

EXAMPLE. A global criterion can be derived from (ii) as follows: assume there exists  $\kappa > 1$  and a continuous function  $\tilde{\theta} : (0, +\infty) \rightarrow [0, 1)$  such that the following global “Weak Asymptotic Confluence” condition holds

$$(WAC)_{\kappa, \tilde{\theta}}^{glob} \equiv \begin{cases} \forall x, y \in \mathbb{R}^d, x \neq y, \\ (b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2 \\ < \frac{1}{2} \left| (\sigma^*(x) - \sigma^*(y)) \frac{x-y}{|x-y|} \right|^2 \left( \left( 1 + \frac{\kappa}{\log(|x-y|^2)} \right) \vee \tilde{\theta}(|x-y|^2) \right). \end{cases}$$

Note that at every couple  $(x, y)$ ,  $x \neq y$ , such that  $(\sigma^*(x) - \sigma^*(y))(x - y) = 0$ , one must have  $(b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2 < 0$ : this is the reason why this criterion *cannot be localized in the neighbourhood of the diagonal* as long as one only relies on Assumption (i).

### 3.4 Local criterion near the diagonal

By considering now Assumption (ii') one can derive a “local” criterion. By local we mean that the weak mean reverting condition is only required in the neighbourhood of the diagonal  $\Delta_{\mathbb{R}^{2d}}$ .

**PROPOSITION 3.2.** *Assume  $b$  and  $\sigma$  are Lipschitz continuous. Let  $\kappa > 1$ . If there exists  $\varepsilon_0 \in (0, e^{-\frac{\kappa}{2}})$*

$$(WAC)_{\kappa, \varepsilon_0}^{logloc} \equiv \begin{cases} (i) \forall x, y \in \mathbb{R}^d, 0 < |x - y| \leq \varepsilon_0, \\ (b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2 \leq \frac{1}{2} \left| (\sigma^*(x) - \sigma^*(y)) \frac{x-y}{|x-y|} \right|^2 \left( 1 + \frac{\kappa}{\log|x-y|^2} \right), \\ (ii) \eta_0 = \inf \left\{ \left| (\sigma^*(x) - \sigma^*(y))(x - y) \right|, |x - y| \geq \varepsilon_0 \right\} > 0, \end{cases}$$

*then the conclusions of Theorem 3.2 hold true (since its assumptions (i)-(ii') are satisfied).*

*Proof.* For every  $u \in (0, +\infty)$

$$\sup_{|x-y|=u} \frac{(b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2}{\left| (\sigma^*(x) - \sigma^*(y)) \frac{x-y}{|x-y|} \right|^2} \leq \theta(u)$$

where

$$\theta(u) = \begin{cases} \theta(u) = ([b]_{\text{Lip}}^2 + \frac{1}{2} [\sigma]_{\text{Lip}}^2) \frac{u^4}{\eta_0^2} & \text{if } u \in [\varepsilon_0, \infty), \\ \theta(u) = \frac{1}{2} \left( 1 + \frac{\kappa}{2 \log u} \right) & \text{if } u \in (0, \varepsilon_0/2) \end{cases}$$



where the second case follows from  $(WAC)_{\kappa, \varepsilon_0}^{logloc}(i)$ . Consequently  $\theta$  can be extended into a continuous function over  $(0, +\infty)$  (by an affine extension between  $\varepsilon_0/2$  and  $\varepsilon_0$ ). Let  $\varepsilon'_0 = \frac{\varepsilon_0}{2} < 1$ . Then, it is clear that for every  $\varepsilon \in (0, \varepsilon'_0]$ ,

$$\int_{\varepsilon}^{\varepsilon'_0} e^{2 \int_v^{\varepsilon'_0} \frac{\theta(w)}{w} dw} dv = K_0 \int_{\varepsilon}^{\varepsilon'_0} \frac{dv}{v |\log v|^\kappa} \longrightarrow \frac{K_0}{(\kappa - 1)(\log(1/\varepsilon'_0))^{\kappa-1}} \text{ as } \varepsilon \rightarrow 0. \quad \square$$

**REMARK 3.7.** • For any matrix  $A \in \mathcal{M}(d, d, \mathbb{R})$ , Schwarz's Inequality trivially implies that

$$\forall u \in \mathbb{R}^d, |u| = 1, |A^*u|^2 \leq \|A\|^2$$

so that the above  $(WAC)_{\kappa, \varepsilon_0}^{loc}(i)$  condition implies the following ‘‘confluence assumption’’ on  $b$ : there exists  $\kappa \in (1, +\infty)$  such that, if  $|x - y| \leq \varepsilon_0$  then

$$(b(x) - b(y)|x - y) \leq \frac{1}{2} \left( \underbrace{\left| (\sigma^*(x) - \sigma^*(y)) \frac{x - y}{|x - y|} \right|^2}_{\leq 0} - \|\sigma(x) - \sigma(y)\|^2 \right) + \underbrace{\frac{\kappa}{2 \log(|x - y|^2)} \left| (\sigma^*(x) - \sigma^*(y)) \frac{x - y}{|x - y|} \right|^2}_{\leq 0}.$$

- The above assumption  $(WAC)_{\kappa, \varepsilon_0}^{logloc}(i)$  is of course satisfied if one replaces the function  $\left(1 + \frac{\kappa}{2 \log(|x - y|^2)}\right)$  by a constant  $\kappa' \in [0, 1)$  on the interval  $(0, \varepsilon_0]$ .
- One can of course carry on the process following the lines of Bertrand's rules: if  $\kappa = 1$ , one may consider an additional term in  $\log |\log |x - y||$  at a power  $\kappa' > 1$ , etc.

### 3.5 Local criterion on compact sets

The conditions  $(WAC)_{\kappa, \tilde{\theta}}^{glob}$  ( $\kappa = 0$ ) are always fulfilled if the global ‘‘Asymptotic Confluence’’ condition:

$$(AC) \equiv \forall x, y \in \mathbb{R}^d, x \neq y, \quad (b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2 < 0 \quad (3.9)$$

is satisfied. One asset of this more stringent assumption is that it can be localized in two ways: first in the neighbourhood of the diagonal like in the above local criterions, but also on compact sets of  $\mathbb{R}^{2d}$ . This naturally leads to a criterion based on the differentials of  $b$  and  $\sigma$  when they exist.

**PROPOSITION 3.3** (Criterion on compact sets). (a) *Assume that*

$$(AC)^{loc^2} \equiv \forall R > 0, \exists \delta_R > 0, \text{ such that } \forall x, y \in B_{|\cdot|}(0; R), 0 < |x - y| \leq \delta_R \implies (b(x) - b(y)|x - y) + \frac{1}{2} \|\sigma(x) - \sigma(y)\|^2 < 0$$

then  $(AC)$  holds.

(b) *If  $b$  and  $\sigma$  are continuously differentiable, then (3.9) holds as soon as*

$$(AC)^{diff} \equiv \forall x \in \mathbb{R}^d, \quad J_b(x) + J_b^*(x) + \sum_{i,j} (\nabla \sigma_{ij})^{\otimes 2}(x) < 0 \text{ in } \mathcal{S}(d, \mathbb{R})$$

where  $J_b(x) = \left[ \frac{\partial b_i}{\partial x_j} \right]_{1 \leq i, j \leq d}$  denotes the Jacobian of  $b(x)$  and  $u^{\otimes 2} = [u_i u_j]_{1 \leq i, j \leq d}$ .

**Proof.** (a) Let  $x, y \in \mathbb{R}^d$  such that  $x \neq y$ . Set  $R = \max(|x|, |y|)$  and

$$x_0 = x, \quad x_i = x + \frac{i}{N}(y - x), \quad i = 1, \dots, N-1, \quad x_N = y$$

where  $|y - x| < N\delta_R$ . Then for every  $i \in \{1, \dots, N\}$ ,  $|x_i| \leq R$  and  $|x_i - x_{i-1}| \leq \delta_R$ . Then

$$\begin{aligned} \|\sigma(x) - \sigma(y)\|^2 &= \left\| \sum_{i=1}^N \sigma(x_i) - \sigma(x_{i-1}) \right\|^2 \leq N \sum_{i=1}^N \|\sigma(x_i) - \sigma(x_{i-1})\|^2 \\ &< -2N \sum_{i=1}^N (b(x_i) - b(x_{i-1}))|x_i - x_{i-1}| = -2 \sum_{i=1}^N (b(x_i) - b(x_{i-1}))|y - x| \\ &< -2(b(y) - b(x))|y - x|. \end{aligned}$$

This shows that  $(AC)^{loc^2}$  implies  $(AC)$ .

(b) Using that for every continuously differentiable function  $g : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $g(y) - g(x) = \int_0^1 (\nabla g(x + t(y-x))|y-x) dt = \int_0^1 (y-x)^* \nabla g(x + t(y-x)) dt$ , one derives that

$$(b(y) - b(x))|y - x| = \int_0^1 (y-x)^* J_b(x + t(y-x))(y-x) dt$$

and

$$\|\sigma(y) - \sigma(x)\|^2 = \sum_{i,j=1}^d \left( \int_0^1 (\nabla \sigma_{ij}(x + t(y-x))|y-x) dt \right)^2.$$

Hölder's Inequality then implies

$$\begin{aligned} (b(y) - b(x))|y - x| + \frac{1}{2} \|\sigma(y) - \sigma(x)\|^2 &\leq \int_0^1 \left( (y-x)^* J_b(x + t(y-x))(y-x) \right. \\ &\quad \left. + \frac{1}{2} \sum_{ij} (y-x)^* (\nabla \sigma_{ij}(x + t(y-x)))^{\otimes 2} (y-x) \right) dt \end{aligned}$$

since  $(\nabla \sigma_{ij}(x + t(y-x))|y-x)^2 = (y-x)^* (\nabla \sigma_{ij}(x + t(y-x)))^{\otimes 2} (y-x)$ . The conclusion follows.  $\square$

- Assume  $\rho^* \rho = I_q$  i.e.  $\rho \in \mathcal{O}(q, \mathbb{R})$ . This case seems of little interest since  $W^{(\rho)} = \rho W$  is still a standard B.M. (think to  $\rho = -1$  when  $d = 1$ ). However most of what precedes remains true by replacing  $\sigma(y)$  by  $\sigma(y)\rho$ , except for the criterion involving differentials.

## 4 Application to the Richardson-Romberg extrapolation for the approximation of invariant distributions

As an application, we investigate in this section the Richardson-Romberg ( $RR$ ) extrapolation for the approximation of invariant measures. Roughly speaking, the aim of a  $RR$  method is generally to improve the order of convergence of an algorithm based on an discretization scheme by cancelling the first order error term induced by the time discretization of the underlying process. However, to be efficient, such a method must be implemented with a control of its variance. We will see that in this context, this control is strongly linked to the uniqueness of the invariant distribution of the duplicated diffusion.

## 4.1 Setting and Background

### 4.1.1 Recursive computation of the invariant distribution of a diffusion: the original procedure

Following [LP02] and a series of papers cited in the introduction, we consider here a sequence of empirical measures  $(\nu_n^\eta(\omega, dx))_{n \geq 1}$  built as follows: let  $(\gamma_n)_{n \geq 1}$  denote a non-increasing sequence of positive *step parameters* satisfying

$$\gamma_n \xrightarrow{n \rightarrow +\infty} 0 \quad \text{and} \quad \Gamma_n = \sum_{k=1}^n \gamma_k \xrightarrow{n \rightarrow +\infty} +\infty.$$

We denote by  $(\bar{X}_n)_{n \geq 0}$  the Euler scheme with step sequence  $(\gamma_n)_{n \geq 1}$  defined by  $\bar{X}_0 = x \in \mathbb{R}^d$  and

$$\bar{X}_{n+1} = \bar{X}_n + \gamma_{n+1} b(\bar{X}_n) + \sqrt{\gamma_{n+1}} \sigma(\bar{X}_n) U_{n+1}$$

where  $(U_n)_{n \geq 1}$  is a sequence of i.i.d. centered  $\mathbb{R}^q$ -valued random vectors such that  $\Sigma_{U_1} = I_q$  defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . The sequence of *weighted empirical measures*  $(\nu_n^\eta(\omega, dx))_{n \geq 1}$  is then defined for every  $n \geq 1$ , by

$$\nu_n^\eta(\omega, f) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{\bar{X}_{k-1}(\omega)}$$

where  $\delta_a$  denotes the Dirac mass at  $a \in \mathbb{R}^d$  and  $(\eta_k)_{k \geq 1}$  is a sequence of positive weights such that  $H_n = \sum_{k=1}^n \eta_k \xrightarrow{n \rightarrow +\infty} +\infty$ . When  $\eta_k = \gamma_k$  which corresponds to the genuine case, we will only write  $\nu_n(\omega, dx)$  instead of  $\nu_n^\eta(\omega, dx)$ . For this sequence, we recall in Proposition 4.4 below in a synthesized form the main convergence results (including rates) of the sequence  $(\nu_n^\eta(\omega, dx))$  to the invariant distribution  $\nu$  of  $(X_t)$ . In this way, we introduce two assumptions:

**(S<sub>a</sub>)** : ( $a > 0$ ) There exists a positive  $\mathcal{C}^2$ -function  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  with

$$\lim_{|x| \rightarrow +\infty} V(x) = +\infty, \quad |\nabla V|^2 \leq CV, \quad \text{and} \quad \sup_{x \in \mathbb{R}^d} \|D^2 V(x)\| < +\infty$$

such that there exist some positive constants  $C_b, \beta$  and  $\alpha$  such that:

$$(i) \quad |b|^2 \leq C_b V^a, \quad \text{Tr}(\sigma\sigma^*)(x) = o(V^a(x)) \quad \text{as } |x| \rightarrow +\infty \quad (ii) \quad (\nabla V|b) \leq \beta - \alpha V^a.$$

This Lyapunov-type assumption is sufficient to ensure the long-time stability of the Euler scheme (in a sense made precise below) as soon as  $a \in (0, 1]$ . Note that the convergence can be obtained under a less restrictive mean-reverting assumption including the case  $a = 0$  (see [Pan06]). The second assumption below is fundamental to establish the rate of convergence of  $(\nu_n^\eta(\omega, f))$  to  $\nu(f)$  for a fixed smooth enough function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ : we assume that  $f$  has a smooth solution to the Poisson equation (see [PV01] for results on this topic).

**(C(f, k))**: There exists a  $\mathcal{C}^k$ -function  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  solution to  $f - \nu(f) = \mathcal{A}g$  such that  $f, g$  and its partial derivatives up to  $k$  are dominated by  $V^r$  ( $r \geq 0$ ):  $|f| \leq CV^r$  and for every  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$  with  $|\alpha| := \alpha_1 + \dots + \alpha_d \in \{0, \dots, k\}$ ,  $|\partial_{x_1^{\alpha_1}, \dots, x_d^{\alpha_d}} g| \leq CV^r$ .

Before recalling the results on  $(\nu_n(\omega, dx))$ , let us introduce further notations. We set

$$\forall r \in \mathbb{N}, \quad \Gamma_n^{(r)} = \sum_{k=1}^n \gamma_k^r$$

and for a smooth enough function  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  and an integer  $r \geq 2$ , we write:

$$D^{(r)}h(x) y_1 \otimes \cdots \otimes y_r = \sum_{(i_1, \dots, i_r) \in \{1, \dots, d\}^r} \partial_{x_{i_1}, \dots, x_{i_r}}^r h(x) y_1^{i_1} \cdots y_r^{i_r}.$$

**PROPOSITION 4.4.** *Assume  $(\mathbf{S}_a)$  holds for an  $a \in (0, 1]$  and  $U_1 \in \cap_{p>0} L^p(\mathbb{P})$ . Assume that  $(\eta_k/\gamma_k)$  is a non-increasing sequence. Then,*

(i) *For every non-increasing sequence  $(\theta_n)_{n \geq 1}$  such that  $\sum_{n \geq 1} \theta_n \gamma_n < +\infty$  and for every  $r > 0$ ,  $\sum_{n \geq 1} \theta_n \gamma_n \mathbb{E}[V^r(\bar{X}_n)] < +\infty$ .*

(ii) *For every  $r > 0$ ,  $\sup_{n \geq 1} \nu_n^\eta(\omega, V^r) < +\infty$  a.s. In particular,  $(\nu_n^\eta(\omega, dx))_{n \geq 1}$  is a.s. tight.*

(iii) *Every weak limit of  $(\nu_n^\eta(\omega, dx))_{n \geq 1}$  is an invariant distribution for  $(X_t)_{t \geq 0}$ . Furthermore, if (SDE) has a unique invariant distribution, say  $\nu$ , then  $\nu_n^\eta(\omega, f) \xrightarrow{n \rightarrow +\infty} \nu(f)$  a.s. for every  $\nu$ -a.s. continuous function  $f$  such that  $|f| \leq CV^r$  for an  $r > 0$ .*

(iv) *(Rate of convergence when  $\eta_k = \gamma_k$ ): Assume that  $\nu$  is unique and that  $\mathbb{E}[U_1^{\otimes 3}] = 0$ . Let  $k \geq 1$  such that  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfies  $(\mathbf{C}(f, \mathbf{k}))$ . Then,*

- *If  $k = 4$  and  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \xrightarrow{n \rightarrow +\infty} 0$ ,*

$$\sqrt{\Gamma_n}(\nu_n(\omega, f) - \nu(f)) \xrightarrow{(\mathbb{R})} \mathcal{N}\left(0; \int |\sigma^* \nabla g|^2 d\nu\right) \quad \text{as } n \rightarrow +\infty.$$

- *If  $k = 5$  and  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \xrightarrow{n \rightarrow +\infty} \tilde{\beta} \in (0, +\infty]$ ,*

$$\triangleright \sqrt{\Gamma_n}(\nu_n(\omega, f) - \nu(f)) \xrightarrow{(\mathbb{R})} \mathcal{N}\left(\tilde{\beta} m_g^{(1)}; \int |\sigma^* \nabla g|^2 d\nu\right) \quad \text{as } n \rightarrow +\infty \quad \text{if } \tilde{\beta} \in (0, +\infty)$$

$$\triangleright \frac{\Gamma_n}{\Gamma_n^{(2)}}(\nu_n(\omega, f) - \nu(f)) \xrightarrow{\text{a.s.}} m_g^{(1)} \quad \text{as } n \rightarrow +\infty \quad \text{if } \tilde{\beta} = +\infty$$

where  $m_g^{(1)} = \int_{\mathbb{R}^d} \varphi_1 d\nu$  with

$$\varphi_1(x) = \frac{1}{2} D^2 g(x) b(x)^{\otimes 2} + \frac{1}{2} \mathbb{E}[D^3 g(x) b(x) (\sigma(x) U_1)^{\otimes 2}] + \frac{1}{24} \mathbb{E}[D^4 g(x) (\sigma(x) U_1)^{\otimes 4}]. \quad (4.10)$$

The first three claims part (i), (ii) and (iii) of the theorem follow from [LP03] whereas the (iv) is derived from [LP02] (see Theorem 10) and [Lem05] (see Theorem V.3), in which the rate of convergence is established for a wide family of weights  $(\eta_k)$ .

Applying (iv) to polynomial steps of the following form:  $\gamma_n = Cn^{-\mu}$ ,  $\mu \in (0, 1]$ , we observe that the optimal (weak) rate is  $n^{-1/3}$  and is attained for  $\mu = 1/3$ . Then

$$\tilde{\beta} = \sqrt{6} C^{\frac{3}{2}} \quad \text{and} \quad \sqrt{\Gamma_n} \sim \sqrt{3C/2} n^{\frac{1}{3}}.$$

so that

$$n^{\frac{1}{3}}(\nu_n(\omega, f) - \nu(f)) \xrightarrow{(\mathbb{R})} \mathcal{N}\left(2C m_g^{(1)}; \frac{2}{3C} \int |\sigma^* \nabla g|^2 d\nu\right).$$

This corresponds to the case where the rate of convergence of the underlying diffusion toward its steady regime ( $\sqrt{\Gamma_n}$  corresponding to  $\sqrt{t}$  in the continuous time setting, see [Bha82] for the CLT for the diffusion itself) and the discretization error are of the same order. From a practical point of view it seems clear that a balance should be made between the asymptotic bias and the asymptotic variance to specify the constant  $C$ . Under

slightly more stringent assumptions we prove that the  $L^2$ -norm of the error  $\nu_n(\omega, f) - \nu(f)$  satisfies

$$\|\nu_n(\omega, f) - \nu(f)\|_{L^2} \sim n^{-\frac{1}{3}} \sqrt{4C^2(m_g^{(1)})^2 + \frac{2}{3C} \int |\sigma^* \nabla g|^2 d\nu}.$$

An optimisation with respect to  $C$  gives the optimal choice  $C = \left( \frac{12 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu}{(m_g^{(1)})^2} \right)^{\frac{1}{3}}$ .

When  $\mu \in (0, 1/3)$ , the step sequence decreases too slowly and the error induced by the time discretization error becomes prominent. That is why we propose below to use an RR extrapolation in order to cancel the first-order term in the time discretization error: in practice this amount to killing the bias  $m_g^{(1)}$  in order to extend the range of application of the rate  $\sqrt{\Gamma_n}$  (which corresponds to the standard weak rate  $\sqrt{t}$  in Bhattacharia's *CLT*) to "slower steps".

#### 4.1.2 The Richardson-Romberg extrapolated algorithm

As mentioned before, the starting idea is to introduce a second Euler scheme with step sequence  $(\tilde{\gamma}_n)_{n \geq 1}$  defined by

$$\forall n \geq 1, \quad \tilde{\gamma}_{2(n-1)} = \tilde{\gamma}_{2n} = \frac{\gamma_n}{2}.$$

As concerns the white noise of both schemes, our aim is to make them consistent in absolute time and correlated (with correlation matrix  $\rho$  satisfying  $I_q - \rho^* \rho \in \mathcal{S}^+(d, \mathbb{R})$ ). To achieve that we proceed as follows.

Let  $(Z_n)_{n \geq 1}$  be a sequence of *i.i.d.*  $\mathbb{R}^q$ -valued random vectors lying in  $\cap_{p>0} L^p(\mathbb{P})$  and satisfying

$$\mathbb{E} Z_1 = 0, \quad \Sigma_{Z_1} = I_q, \quad \mathbb{E}[Z_1^{\otimes 3}] = \mathbb{E}[Z_1^{\otimes 5}] = 0.$$

Then we devise from this sequence the white noise sequence  $(U_n)_{n \geq 1}$  of the "original" Euler scheme with step  $(\gamma_n)_{n \geq 1}$  by setting

$$\forall n \geq 1, \quad U_n = \frac{1}{\sqrt{2}} (Z_{2n-1} + Z_{2n}). \quad (4.11)$$

The white noise sequence for the second Euler scheme (with step  $(\tilde{\gamma}_n)_{n \geq 1}$ ), denoted  $Z^{(\rho)}$  is defined as follows:

$$Z_n^{(\rho)} = \rho^* Z_n + T(\rho) V_n, \quad n \geq 1, \quad (4.12)$$

where  $(V_n)_{n \geq 1}$  is also a sequence of *i.i.d.* centered random variables in  $\mathbb{R}^q$  with moments of any order satisfying  $\Sigma_{V_1} = I_q$  and  $\mathbb{E}[V_1^{\otimes 3}] = \mathbb{E}[V_1^{\otimes 5}] = 0$ , *independent* of  $(Z_n)_{n \geq 1}$  and  $T_q(\rho)$  is a solution to the equation

$$T_q(\rho) T_q(\rho)^* = I_q - \rho^* \rho \in \mathcal{S}^+(d, \mathbb{R}).$$

( $T_q(\rho)$  can be chosen either as the commuting symmetric square root of  $I_q - \rho^* \rho$  or its Choleski transform). Note that  $(Z_n^{(\rho)})_{n \geq 1}$  is built in so that it satisfies

$$\Sigma_{Z_n^{(\rho)}} = I_q \quad \text{and} \quad \text{Cov}(Z_n, Z_n^{(\rho)}) = \rho.$$

Then the Euler scheme with step  $\tilde{\gamma}_n$  and consistent  $\rho$ -correlated white noise  $(Z_n^{(\rho)})_{n \geq 1}$ , denoted  $(\bar{Y}_n^{(\rho)})_{n \geq 1}$  from now on, is defined by:

$$\bar{Y}_{n+1}^{(\rho)} = \bar{Y}_n^{(\rho)} + \tilde{\gamma}_n b(\bar{Y}_n^{(\rho)}) + \sqrt{\tilde{\gamma}_n} \sigma(\bar{Y}_n^{(\rho)}) Z_{n+1}^{(\rho)}, \quad n \geq 1, \quad \bar{Y}_0 = y.$$

Also note that  $(\bar{X}_n, \bar{Y}_{2n}^{(\rho)})$  is an Euler scheme at time  $\Gamma_n$  of the duplicated diffusion  $(X_t, X_t^{(\rho)})_{t \geq 0}$ .

For numerical purpose, one usually specifies the independent i.i.d. sequences  $(Z_n)_{n \geq 1}$  and  $(V_n)_{n \geq 1}$  as normally distributed so that they can be considered as the normalized increments of two independent Brownian motions  $W$  and  $\widetilde{W}$  *i.e.*

$$Z_n = \frac{W_{\tilde{\Gamma}_n} - W_{\tilde{\Gamma}_{n-1}}}{\sqrt{\tilde{\gamma}_n}} \quad \text{and} \quad V_n = \frac{\widetilde{W}_{\tilde{\Gamma}_n} - \widetilde{W}_{\tilde{\Gamma}_{n-1}}}{\sqrt{\tilde{\gamma}_n}}, \quad n \geq 1.$$

Note that in this case,  $(U_n)$  is also a sequence of  $\mathcal{N}(0, I_q)$ -random variables. This implies in particular that

$$\mathbb{E}[U_1^{\otimes 4}] = \mathbb{E}[Z_1^{\otimes 4}] \quad \text{and} \quad \mathbb{E}[U_1^{\otimes 6}] = \mathbb{E}[Z_1^{\otimes 6}]. \quad (4.13)$$

Since these properties simplify the result, we will assume them in the sequel of this section (see Remark 4.8 for extensions).

We denote  $(\nu_n^{\eta, (\rho)}(\omega, dx))_{n \geq 1}$  the sequence of empirical measures related to  $(\bar{Y}_n^{(\rho)}(\omega))_{n \geq 1}$  (in which the weights are adapted accordingly:  $\eta_1/2, \eta_1/2, \eta_2/2, \eta_2/2, \eta_3/2, \dots$ ). The empirical measure  $(\bar{\nu}_n^{\eta, (\rho)}(\omega, dx))_{n \geq 1}$  associated to the Richardson-Romberg extrapolation is defined by

$$\begin{aligned} \nu_n^{\eta, (\rho)}(\omega, f) &= \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{2} \left( f(\bar{Y}_{2(k-1)}^{(\rho)}(\omega)) + f(\bar{Y}_{2k-1}^{(\rho)}(\omega)) \right) \\ \bar{\nu}_n^{\eta, (\rho)}(\omega, f) &= (2\nu_n^{\eta, (\rho)} - \nu_n^{\eta}(\omega, f)) \\ &= \frac{1}{H_n} \sum_{k=1}^n \eta_k \left( f(\bar{Y}_{2(k-1)}^{(\rho)}(\omega)) + f(\bar{Y}_{2k-1}^{(\rho)}(\omega)) - f(\bar{X}_k(\omega)) \right). \end{aligned}$$

Under the assumptions of Proposition 4.4, it is clear that  $\bar{\nu}_n^{\eta, (\rho)}(\omega, dx) \xrightarrow{n \rightarrow +\infty} \nu(dx)$  *a.s.*

Thus, in the next section, we propose to evaluate the effects of the Richardson-Romberg extrapolation on the rate of convergence of the procedure and to explain why the uniqueness of the invariant distribution of the duplicated diffusion plays an important role in this problem.

## 4.2 Rate of convergence of the extrapolated procedure

Throughout this section we assume that  $\eta_k = \gamma_k$  and so we will write  $\nu_n, \nu_n^{(\rho)}$  and  $\bar{\nu}_n^{(\rho)}$  instead of  $\nu_n^{\eta}, \nu_n^{\eta, (\rho)}$  and  $\bar{\nu}_n^{\eta, (\rho)}$  respectively. We also set  $(D^3 g_{i, \dots})_{i=1}^d = D^2(\nabla \cdot)g$  in order that the notation  $\text{Tr}(\sigma^* D^2(\nabla \cdot)g\sigma)$  stands for the *vector* of  $\mathbb{R}^d$  defined by  $\text{Tr}(\sigma^* D^2(\nabla \cdot)g\sigma) = (\text{Tr}(\sigma^* D^2(\partial_{x_i} g)\sigma))_{i=1}^d$ . For a fixed matrix  $\rho$ , the main result about the *RR* extrapolation is Theorem 4.3 below. At this stage, we do not discuss the choice of the correlation  $\rho$  in this result. This point is tackled in Proposition 4.5 in which we will see that the optimal choice to reduce the asymptotic variance is attained with  $\rho = I_q$  as soon as  $\nu_\Delta$  is the

unique invariant distribution of the associated duplicated diffusion. This emphasizes the importance of the question of the uniqueness of the invariant distribution in this pathologic case studied in the previous part of the paper.

**THEOREM 4.3.** *Assume  $(\mathbf{S}_a)$  holds for an  $a \in (0, 1]$ . Assume that  $(X_t, X_t^{(\rho)})_{t \geq 0}$  admits a unique invariant distribution  $\mu^{(\rho)}$  (with marginals  $\nu$ ). Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be a function satisfying  $(\mathbf{C}(f, \mathbf{7}))$  and such that  $\varphi_1$  defined by (4.10) satisfies  $(\mathbf{C}(\varphi_1, \mathbf{5}))$  with a solution to the Poisson equation denoted by  $g_{\varphi_1}$ . Then,*

- If  $\frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \xrightarrow{n \rightarrow +\infty} 0$ ,

$$\sqrt{\Gamma_n} \left( \nu_n^{(\rho)}(\omega, f) - \nu(f) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0; \hat{\sigma}_\rho^2)$$

where

$$\hat{\sigma}_\rho^2 = 5 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu - 4 \int_{\mathbb{R}^{2d}} ((\sigma^* \nabla g)(x) | \rho(\sigma^* \nabla g)(y)) \mu^{(\rho)}(dx, dy). \quad (4.14)$$

- If  $\frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \xrightarrow{n \rightarrow +\infty} \tilde{\beta} \in (0, +\infty]$ , then

$$\begin{aligned} \sqrt{\Gamma_n} (\nu_n^{(\rho)}(\omega, f) - \nu(f)) &\xrightarrow{(\mathbb{R})} \mathcal{N}(\tilde{\beta} m_g^{(2)}; \hat{\sigma}_\rho^2) \quad \text{as } n \rightarrow +\infty && \text{if } \tilde{\beta} \in (0, +\infty), \\ \frac{\Gamma_n}{\Gamma_n^{(3)}} (\nu_n^{(\rho)}(\omega, f) - \nu(f)) &\xrightarrow{\mathbb{P}} m_g^{(2)} \quad \text{as } n \rightarrow +\infty && \text{if } \tilde{\beta} = +\infty, \end{aligned}$$

where  $m_g^{(2)} = \frac{1}{2} \left( m_{g_{\varphi_1}} + \int_{\mathbb{R}^d} \varphi_2 d\nu \right)$  with

$$\varphi_2(x) = \sum_{k=3}^6 \frac{C_k^{2(k-3)}}{k!} \mathbb{E} [D^k g(x) b(x)^{\otimes(6-k)} (\sigma(x) U_1)^{\otimes 2(k-3)}]. \quad (4.15)$$

**REMARK 4.8.**  $\triangleright$  We recall that the result is stated under the assumption that the increments are normally distributed or more precisely under Assumption (4.13). When this additional assumption fails (think for instance to  $Z_1 \sim (\frac{1}{2}(\delta_{-1} + \delta_1))^{\otimes q}$ ), the result is remains true except for the value of  $m_g^{(2)}$  which becomes more complicated since it also depends on  $\mathbb{E}[Z_1^{\otimes \ell}]$ ,  $\ell = 4$  and  $6$ ).

$\triangleright$  This result extends readily to general weights sequences  $(\eta_n)_{n \geq 1}$ . Some technical conditions appear on the choice of weights but these conditions are natural and not restrictive (see [Lem05]). In particular we can always consider the choice  $\eta_n = 1$  for which we obtain the following result: if  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n^{(-1)}}} \xrightarrow{n \rightarrow +\infty} \tilde{\beta} \in (0, +\infty)$ , then

$$\frac{n}{\sqrt{\Gamma_n^{(-1)}}} (\nu_n^{(\rho)}(\omega, f) - \nu(f)) \xrightarrow{(\mathbb{R})} \mathcal{N}(\tilde{\beta} m_g^{(2)}; \hat{\sigma}_\rho^2) \quad \text{as } n \rightarrow +\infty.$$

$\triangleright$  POLYNOMIAL STEPS. Let  $\gamma_n = Cn^{-\mu}$ ,  $\mu \in (0, 1]$ . If  $\mu > \frac{1}{3}$ ,  $\Gamma_n^{(3)} \rightarrow \Gamma_\infty^{(3)} < +\infty$  so that  $\frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \rightarrow 0$  as  $n \rightarrow +\infty$ . If  $\mu < \frac{1}{3}$ ,  $\frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \asymp n^{\frac{1-5\mu}{2}}$  (and if  $\mu = \frac{1}{3}$ ,  $\frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \asymp \frac{\log n}{\sqrt{n}}$ ). Consequently

$$\frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \rightarrow 0 \iff \mu > \frac{1}{5}, \quad \frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \rightarrow +\infty \iff \mu < \frac{1}{5} \quad \text{and} \quad \frac{\Gamma_n^{(3)}}{\sqrt{\Gamma_n}} \rightarrow \tilde{\beta} \in (0, +\infty) \iff \mu = \frac{1}{5}.$$

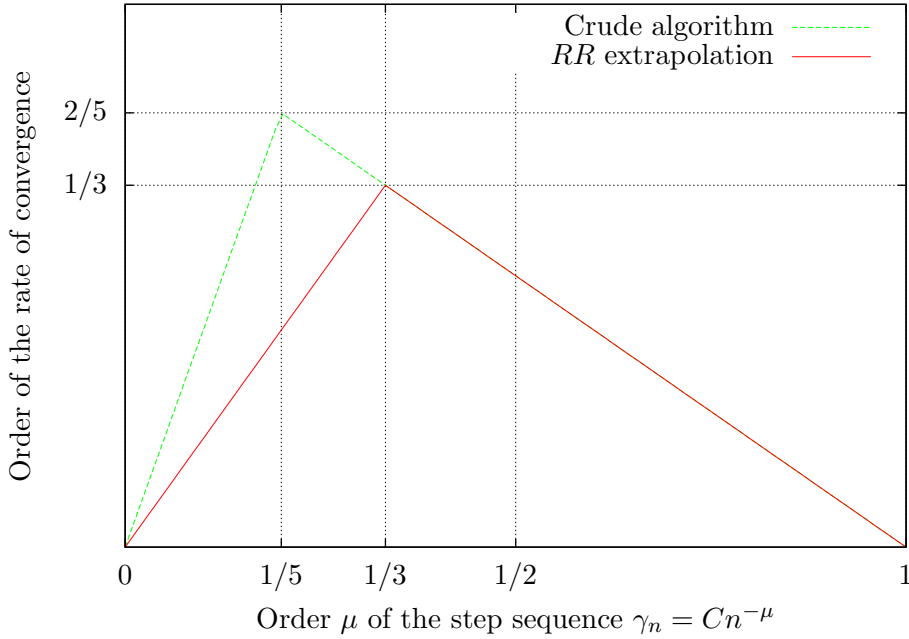
When  $\mu = \frac{1}{5}$ ,  $\tilde{\beta} = C^{\frac{5}{2}}\sqrt{5}$  and  $\sqrt{\Gamma_n} \sim \frac{\sqrt{5}C}{2}n^{\frac{2}{5}}$ .

As a consequence, if  $\gamma_n = \eta_n = Cn^{-\frac{1}{5}}$ ,

$$n^{\frac{2}{5}}(\nu_n^{(\rho)}(\omega, f) - \nu(f)) \xrightarrow{(\mathbb{R})} \mathcal{N}\left(2C^2 m_g^{(2)}; \frac{4}{5} \frac{\hat{\sigma}_\rho^2}{C}\right).$$

We switch from a weak rate  $n^{\frac{1}{3}}$  to  $n^{\frac{2}{5}}$  i.e. a “gain” of  $n^{\frac{1}{15}}$  (see figure below). The second noticeable fact is that the bias is now significantly more sensitive to the constant  $C$  than in the standard setting. If we minimize the  $L^2$ -norm of the error  $\nu_n^{(\rho)}(\omega, f) - \nu(f)$  we obtain the optimal choice of  $C$  as a function of both bias and standard deviation, precisely

$$C = \left(\frac{\hat{\sigma}_\rho^2}{20(m_q^{(2)})^2}\right)^{\frac{1}{5}}.$$



#### 4.2.1 Optimal choice of $\rho$ and uniqueness of $\mu^{(I_d)}$

**PROPOSITION 4.5.** *Let  $\rho$  be an admissible correlation matrix i.e. such that  $\rho^*\rho \leq I_q$ . Assume that the duplicated diffusion  $(X, X^{(\rho)})$  has a unique invariant distribution  $\mu^{(\rho)}$  (so that if  $\rho = I_q$ ,  $\mu^{(I_q)} = \nu_\Delta$ ).*

(a)  $\hat{\sigma}_\rho^2 \geq \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu.$

(b) If  $\rho = 0$  then  $\hat{\sigma}_\rho^2 = 5 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu.$

(c) If  $\rho = I_q$ ,  $\hat{\sigma}_\rho^2 = \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu.$

*Proof.* Claims (b) and (c) being obvious thanks to (4.14), we only prove (a). Keeping in mind that both marginals  $\mu^{(\rho)}(\mathbb{R}^d \times dy)$  and  $\mu^{(\rho)}(dx \times \mathbb{R}^d)$  are equal to  $\nu$ , one derives



thanks to Schwarz's Inequality (once on  $\mathbb{R}^d$  and once on  $L^2(\mu)$ ) from the expression (4.14) of the asymptotic variance  $\widehat{\sigma}_\rho^2$  that

$$\begin{aligned}\widehat{\sigma}_\rho^2 &\geq 5 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu - 4 \left[ \int_{\mathbb{R}^{2d}} |\sigma^* \nabla g|^2(x) \mu^{(\rho)}(dx, dy) \right]^{\frac{1}{2}} \left[ \int_{\mathbb{R}^{2d}} |\rho \sigma^* \nabla g|^2(y) \mu^{(\rho)}(dx, dy) \right]^{\frac{1}{2}} \\ &= 5 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu - 4 \left[ \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2(x) \nu(dx) \right]^{\frac{1}{2}} \left[ \int_{\mathbb{R}^d} |\rho \sigma^* \nabla g|^2(y) \nu(dy) \right]^{\frac{1}{2}} \\ &\geq 5 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu - 4 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu = \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu\end{aligned}$$

where we used in the last inequality that  $|\rho u|^2 \leq |u|^2$ .

The previous result says that the structural asymptotic variance of the  $RR$  estimator is always greater than that of the standard estimator but can be equal if the Brownian motions are equal. This condition is in fact almost necessary. Actually, thanks to the Pythagorean identity,

$$\begin{aligned}\sigma_\rho^2 &= 5 \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu + 2 \int_{\mathbb{R}^{2d}} |\sigma^* \nabla g(x) - \rho \sigma^* \nabla g(y)|^2 \mu^{(\rho)}(dx, dy) \\ &\quad - 2 \int_{\mathbb{R}^{2d}} |\sigma^* \nabla g(x)|^2 \nu(dx) - 2 \int_{\mathbb{R}^{2d}} |\rho \sigma^* \nabla g(y)|^2 \nu(dy).\end{aligned}$$

Then, since  $\rho^* \rho \leq I_q$ , a necessary condition to obtain  $\sigma_\rho^2 = \int_{\mathbb{R}^d} |\sigma^* \nabla g|^2 d\nu$  is

$$|\rho \sigma^* \nabla g(y)| = |\sigma^* \nabla g(y)| \quad \nu(dy)\text{-a.e.}$$

When  $\rho^* \rho < I_q$ , this equality can not hold except if  $\sigma^* \nabla g(y) = 0 \quad \nu(dy)\text{-a.e.}$

### 4.3 Proof of Theorem 4.3

#### 4.3.1 Preliminaries

Without loss of generality, we assume that  $f$  satisfies  $\nu(f) = 0$  so that  $f = \mathcal{A}g$  under  $(\mathbf{C}(\mathbf{f}, \mathbf{k}))$ . We denote by  $\gamma^{(r)}$  the sequence defined by  $\gamma_k^{(r)} = \gamma_k^r$ .

**LEMMA 4.1.** *Assume that  $f$  satisfies  $(\mathbf{C}(\mathbf{f}, \mathbf{7}))$  and denote by  $g$  the solution to the Poisson equation  $\mathcal{A}g = f$ . Then,*

$$\begin{aligned}\Gamma_n \bar{\nu}_n^{(\rho)}(\omega, f) &= 2(g(\bar{Y}_{2n}) - g(\bar{Y}_0)) - (g(\bar{X}_n) - g(\bar{X}_0)) - \sum_{k=1}^n \sqrt{\gamma_k} (\sqrt{2} \Delta M_k^{(2)} - \Delta M_k^{(1)}) \quad (4.16) \\ &\quad - \mathcal{E}_n^1 - \mathcal{E}_n^2 + N_n + R_n \quad (4.17)\end{aligned}$$

where

$$\begin{aligned}\Delta M_k^{(1)} &= (\nabla g(\bar{X}_{k-1}) | \sigma(\bar{X}_{k-1}) U_k), \\ \Delta M_k^{(2)} &= (\nabla g(\bar{Y}_{2(k-1)}) | \sigma(\bar{Y}_{2(k-1)}) Z_{2k-1}) + (\nabla g(\bar{Y}_{2k-1}) | \sigma(\bar{Y}_{2k-1}) Z_{2k}), \\ \mathcal{E}_n^1 &= 2 \sum_{k=1}^n \left( \frac{\gamma_k}{2} \right)^2 (\varphi_1(\bar{Y}_{2(k-1)}) + \mathbb{E}[\varphi_1(\bar{Y}_{2k-1}) | \mathcal{F}_{k-1}]) - \sum_{k=1}^n \gamma_k^2 \varphi_1(\bar{X}_{k-1}), \\ \mathcal{E}_n^2 &= 2 \sum_{k=1}^n \left( \frac{\gamma_k}{2} \right)^3 (\varphi_2(\bar{Y}_{2(k-1)}) + \mathbb{E}[\varphi_2(\bar{Y}_{2k-1}) | \mathcal{F}_{k-1}]) - \sum_{k=1}^n \gamma_k^3 \varphi_2(\bar{X}_{k-1})\end{aligned}$$

with  $\varphi_1$  and  $\varphi_2$  defined by (4.10) and (4.15),

$(N_n)$  is defined by

$$N_n = \sum_{k=1}^n 2 \left( \Delta N(\bar{Y}_{2(k-1)}, Z_{2k-1}, \frac{\gamma_k}{2}) + \Delta N(\bar{Y}_{2k-1}, Z_{2k}, \frac{\gamma_k}{2}) \right) - \Delta N(\bar{X}_{k-1}, Z_{2k-1}, \frac{\gamma_k}{2})$$

where  $\Delta N(x, U, \gamma) = H(x, U, \gamma) - \mathbb{E}_x[H(x, U, \gamma)]$  and

$$\begin{aligned} H(x, U, \gamma) &= \frac{\gamma}{2} D^2 g(x) (\sigma(x)U)^{\otimes 2} + \frac{1}{6} \sum_{\ell=0}^2 C_3^{3-\ell} \gamma^{\frac{\ell+3}{2}} D^3 g(x) b(x)^{\otimes \ell} (\sigma(x)U)^{\otimes (3-\ell)} \\ &+ \frac{1}{24} \sum_{\ell=0}^1 \gamma^{\frac{\ell+4}{2}} C_4^{4-\ell} D^4 g(x) b(x)^{\otimes \ell} (\sigma(x)U)^{\otimes (4-\ell)} + \gamma^3 \sum_{\ell=4}^6 \frac{C_\ell^{6-\ell}}{\ell!} D^\ell g(x) b(x)^{\otimes (6-\ell)} (\sigma(x)U)^{\otimes \frac{\ell}{2}}. \end{aligned}$$

Finally, if  $(\mathbf{S}_a)$  holds, the sequence  $(R_n)_{n \geq 1}$  satisfies the following property: there exists  $r > 0$  such that, a.s., for every  $n \geq 1$ ,

$$\mathbb{E}[|\Delta R_n| | \mathcal{F}_{n-1}] \leq C \gamma_n^{\frac{7}{2}} (V^r(\bar{X}_{n-1}) + V^r(\bar{Y}_{2(n-1)}) + V^r(\bar{Y}_{2n-1})) \quad (4.18)$$

where  $\Delta R_n = R_n - R_{n-1}$ .

**REMARK 4.9.** The above decomposition is built as follows: the second term of (4.16) is the main martingale component of the decomposition whereas  $\mathcal{E}_{n,1}$  contains the first order discretization error. Thanks to the Richardson-Romberg extrapolation,  $\mathcal{E}_{n,1}$  is in fact negligible when  $n \rightarrow +\infty$ . When the step sequence decreases fast (Theorem 4.3(i)), the rate of convergence is ruled by the main martingale component. In Theorem 4.3(ii), the rate is ruled by  $\mathcal{E}_{n,1}$  and  $\mathcal{E}_{n,2}$ . Finally,  $N_n$  contains all the negligible martingale terms.

*Proof.* Owing to  $(\mathbf{C}(\mathbf{f}, \mathbf{7}))$ , to the Taylor formula and to the fact that  $\mathbb{E}[D^2(x)(\sigma(x)U_1)^{\otimes 2}] = \text{Tr}(\sigma^*(x)D^2g(x)\sigma(x))$ , we have

$$g(\bar{X}_k) = g(\bar{X}_{k-1}) + \gamma_k \mathcal{A}g(\bar{X}_{k-1}) + \sqrt{\gamma_k} \Delta M_{k,1} \quad (4.19)$$

$$+ \frac{1}{2} (D^2(\bar{X}_{k-1})(\sigma(\bar{X}_{k-1})U_k)^{\otimes 2} - \mathbb{E}[D^2(\bar{X}_{k-1})(\sigma(\bar{X}_{k-1})U_k)^{\otimes 2} | \mathcal{F}_{k-1}]) \quad (4.20)$$

$$+ \sum_{l=3}^5 D^l g(\bar{X}_{k-1}) (\gamma_k b(\bar{X}_{k-1}) + \sqrt{\gamma_k} \sigma(\bar{X}_{k-1})U_k)^{\otimes l} \quad (4.21)$$

$$+ D^7 g(\xi_k) (\gamma_k b(\bar{X}_{k-1}) + \sqrt{\gamma_k} \sigma(\bar{X}_{k-1})U_k)^{\otimes 7} \quad (4.22)$$

where  $\xi_k \in [\bar{X}_{k-1}, \bar{X}_k]$ . The fact that  $|\nabla V|^2 \leq CV$  implies that  $\sqrt{V}$  is a Lipschitz continuous function with Lipschitz constant denoted by  $[\sqrt{V}]_{\text{Lip}}$ . Then, setting  $\|D^7 g(x)\| = \sup_{|\alpha|=7} |\partial_\alpha g(x)|$  and using Assumption  $(\mathbf{C}(\mathbf{f}, \mathbf{7}))$ , we have

$$\|D^7 g(\xi_k)\| \leq C(\sqrt{V}(\xi_k))^{2r} \leq C(\sqrt{V}(\bar{X}_{k-1}) + [\sqrt{V}]_{\text{Lip}} |\Delta \bar{X}_k|)^{2r} \quad (4.23)$$

where  $\Delta \bar{X}_k = \gamma_k b(\bar{X}_{k-1}) + \sqrt{\gamma_k} \sigma(\bar{X}_{k-1})U_k$ . Then, owing to the elementary inequality  $|a + b|^p \leq c_p(|a|^p + |b|^p)$  and to Assumption  $(\mathbf{S}_a)$ , it follows that there exists  $r > 0$  such that

$$\mathbb{E}[|\Delta R_n| | \mathcal{F}_{k-1}] \leq C \gamma_k^{\frac{7}{2}} V^r(\bar{X}_{k-1}).$$

Then we plug this control into the above Taylor expansion and to compensate the terms of (4.21) when necessary. An appropriate (tedious) grouping of the terms yields:

$$\begin{aligned}\gamma_k \mathcal{A}g(\bar{X}_{k-1}) &= g(\bar{X}_k) - g(\bar{X}_{k-1}) - \sqrt{\gamma_k} \Delta M_{k,1} \\ &\quad - \gamma_k^2 \varphi_1(\bar{X}_{k-1}) - \gamma_k^3 \varphi_2(\bar{X}_{k-1}) - \Delta N(\bar{X}_{k-1}, U_k, \gamma_k) - \Delta R_n\end{aligned}$$

where  $R_{n,2}$  satisfies (4.18). Making the same development for  $\mathcal{A}g(\bar{Y}_{2(k-1)})$  and for  $\mathcal{A}g(\bar{Y}_{2k-1})$  and summing over  $n$  yield the announced result.  $\square$

**LEMMA 4.2.** *Let  $a \in (0, 1]$  such that  $(\mathbf{S}_a)$  holds. Assume that  $(X_t, X_t^{(\rho)})_{t \geq 0}$  admits a unique invariant distribution  $\mu^{(\rho)}$ . Let  $g$  be a  $\mathcal{C}^1$ -function such that  $|\nabla g| \leq CV^r$  where  $r \in \mathbb{R}_+$ . Then,*

$$\frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \sqrt{\gamma_k} (\sqrt{2} \Delta M_k^{(2)} - \Delta M_k^{(1)}) \xrightarrow{n \rightarrow +\infty} \hat{\sigma}_\rho^2.$$

*Proof.* Let  $\{\xi_{k,n}, k = 1, \dots, n, n \geq 1\}$  be the triangular array of  $(\mathcal{F}_k)$ -martingale increments defined by

$$\xi_{k,n} = \sqrt{\frac{\gamma_k}{\Gamma_n}} (\sqrt{2} \Delta M_k^{(2)} - \Delta M_k^{(1)}).$$

Let us show that

$$\sum_{k=1}^n \mathbb{E}[|\xi_{k,n}|^2 | \mathcal{F}_{k-1}] \xrightarrow{n \rightarrow +\infty} \hat{\sigma}_\rho^2.$$

First, using that  $\Sigma_{U_1} = I_q$ , we obtain that for every  $k \geq 1$ ,

$$\mathbb{E}[|\Delta M_k^{(1)}|^2 | \mathcal{F}_{k-1}] = |\sigma^* \nabla g(\bar{X}_{k-1})|^2.$$

Since  $x \mapsto |\sigma^* \nabla g|^2(x)$  is a continuous function such that  $|\sigma^* \nabla g|^2 \leq CV^r$  for a positive  $r$ , it follows from Proposition (4.4) that

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathbb{E}[|\Delta M_k^{(1)}|^2 | \mathcal{F}_{k-1}] \xrightarrow{n \rightarrow +\infty} \int |\sigma^* \nabla g|^2(x) \nu(dx). \quad (4.24)$$

Similarly,

$$\mathbb{E}[|\Delta M_k^{(2)}|^2 | \mathcal{F}_{k-1}] = |\sigma^* \nabla g(\bar{Y}_{2(k-1)})|^2 + \mathbb{E}[|\sigma^* \nabla g(\bar{Y}_{2k-1})|^2 | \mathcal{F}_{k-1}].$$

It follows that

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathbb{E}[|\Delta M_k^{(2)}|^2 | \mathcal{F}_{k-1}] = 2\nu_n^{(\rho)}(\omega, |\sigma^* \nabla g|^2) - \frac{1}{\Gamma_n} \sum_{k=1}^n \zeta_k \quad (4.25)$$

where  $(\zeta_k)$  is a sequence of  $(\mathcal{F}_k)$ -martingale increments defined by

$$\zeta_k = \gamma_k (|\sigma^* \nabla g(\bar{Y}_{2k-1})|^2 - \mathbb{E}[|\sigma^* \nabla g(\bar{Y}_{2k-1})|^2 | \mathcal{F}_{k-1}]).$$

Using that  $|\sigma^* \nabla g|^2 \leq CV^r$  for a positive real number  $r$ , we obtain by similar arguments to those used in (4.23) that  $\mathbb{E}[|\zeta_k|^2 | \mathcal{F}_{k-1}] \leq CV^{2r}(\bar{Y}_{2(k-1)})$ . We derive from Proposition 4.4(i) applied with  $\theta_k = \frac{1}{\Gamma_k^2}$  that

$$\sum_{k=1}^{+\infty} \mathbb{E}\left[\left|\frac{\zeta_k}{\Gamma_k}\right|^2 \middle| \mathcal{F}_{k-1}\right] \leq C\gamma_1 \sum_{k=1}^{+\infty} \frac{\gamma_k}{\Gamma_k^2} V^{2r}(\bar{Y}_{2(k-1)}) < +\infty$$

since

$$\sum_{k \geq 1} \frac{\gamma_k}{\Gamma_k^2} \leq 1 + \sum_{k=2}^{+\infty} \int_{\Gamma_{k-1}}^{\Gamma_k} \frac{ds}{s^2} \leq 1 + \int_{\Gamma_1}^{+\infty} \frac{ds}{s^2} < +\infty.$$

As a consequence  $(\sum_{k=1}^n \frac{\zeta_k}{\Gamma_k})_{n \geq 1}$  is a convergent martingale and the Kronecker Lemma then implies that  $\frac{1}{\Gamma_n} \sum_{k=1}^n \zeta_k \xrightarrow{n \rightarrow +\infty} 0$  *a.s.* Thus, we deduce from (4.25) combined with Proposition 4.4 that

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathbb{E}[|\Delta M_k^{(2)}|^2 | \mathcal{F}_{k-1}] \xrightarrow{n \rightarrow +\infty} 2\nu(|\sigma^* \nabla g|^2) \quad a.s. \quad (4.26)$$

Finally, we have to manage the cross-product: keeping in mind the construction of the noises of the Euler schemes (see (4.11) and (4.12), we have:

$$\begin{aligned} \sqrt{2} \mathbb{E}[\Delta M_k^{(1)} \Delta M_{k,2} | \mathcal{F}_{k-1}] &= ((\sigma^* \nabla g)(\bar{X}_{k-1}) | \rho(\sigma^* \nabla g)(\bar{Y}_{2(k-1)})) \\ &+ ((\sigma^* \nabla g)(\bar{X}_{k-1}) | \rho(\sigma^* \nabla g)(\bar{Y}_{2k-1})) - \gamma_k^{-1} \zeta_k^{(2)} \end{aligned}$$

where

$$\zeta_k^{(2)} = \gamma_k \left( ((\sigma^* \nabla g)(\bar{X}_{k-1}) | \rho(\sigma^* \nabla g)(\bar{Y}_{2k-1})) - \mathbb{E}[\left( (\sigma^* \nabla g)(\bar{X}_{k-1}) | \rho(\sigma^* \nabla g)(\bar{Y}_{2k-1}) \right) | \mathcal{F}_{k-1}] \right)$$

so that

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \sqrt{2} \mathbb{E}[\Delta M_k^{(1)} \Delta M_{k,2} | \mathcal{F}_{k-1}] = \mu_n^{(1)}(\psi) + \mu_n^{(2)}(\psi) - \frac{1}{\Gamma_n} \sum_{k=1}^n \zeta_k^{(2)} \quad (4.27)$$

where  $\psi : \mathbb{R}^{2d} \rightarrow \mathbb{R}$  is defined by  $\psi(x, y) = (\sigma^* \nabla g(x) | \rho(\sigma^* \nabla g)(y))$  and for every Borel function  $f : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ ,

$$\mu_n^{(1)}(f) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(\bar{X}_{k-1}, \bar{Y}_{2(k-1)}) \quad \text{and} \quad \mu_n^{(2)}(f) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(\bar{X}_{k-1}, \bar{Y}_{2k-1}).$$

By straightforward adaptations of the proof of Proposition 4.4, we can show that if  $(X_t, X_t^{(\rho)})$  has a unique invariant distribution  $\mu^{(\rho)}$  then, for every continuous function  $f$  such that  $f \leq CV^r$  with  $r > 0$ ,

$$\mu_n^{(i)}(\omega, f) \xrightarrow{n \rightarrow +\infty} \mu^{(\rho)}(f) \quad a.s. \quad \text{with } i = 1, 2.$$

As a consequence,  $\mu_n^{(1)}(\psi) + \mu_n^{(2)}(\psi) \xrightarrow{n \rightarrow +\infty} 2\mu(\psi)$  *a.s.* Finally, by martingale arguments similar to those used for  $(\zeta_k)$ , one checks that  $\Gamma_n^{-1} \sum_{k=1}^n \zeta_k^{(2)} \xrightarrow{n \rightarrow +\infty} 0$  *a.s.* Thus, by (4.24), (4.26) and (4.27), we obtain that

$$\sum_{k=1}^n \mathbb{E}[|\xi_{k,n}|^2 | \mathcal{F}_{k-1}] \xrightarrow{n \rightarrow +\infty} 5\nu(|\sigma^* \nabla g|^2) - 4\mu(\psi) = \hat{\sigma}_\rho^2.$$

Then, the result follows from the CLT for arrays of martingale increments provided that a Lindeberg-type condition is satisfied (see [HH80], Corollary 3.1). To be precise, it is enough to prove that there exists  $\delta > 0$  such that

$$\sum_{k=1}^n \mathbb{E}[|\xi_{k,n}|^{2+\delta} | \mathcal{F}_{k-1}] \xrightarrow{n \rightarrow +\infty} 0 \quad a.s. \quad (4.28)$$

Using Assumption  $(\mathbf{S}_a)$  and the fact  $|\nabla g| \leq CV^r$  ( $r > 0$ ), one can check that there exists  $r > 0$  such that

$$\mathbb{E}[|\xi_{k,n}|^{2+\delta} | \mathcal{F}_{k-1}] \leq C \frac{\gamma_k^{1+\delta}}{\Gamma_n^{1+\delta}} (V^r(\bar{X}_{k-1}) + V^r(\bar{Y}_{2k-1}) + V^r(\bar{Y}_{2(k-1)})).$$

Thus,

$$\sum_{k=1}^n \mathbb{E}[|\xi_{k,n}|^{2+\delta} | \mathcal{F}_{k-1}] \leq C \frac{\Gamma_n^{(1+\delta)}}{\Gamma_n^{1+\delta}} \left( \nu_n^{\gamma^{(1+\delta)}}(V^r) + \nu_n^{\gamma^{(1+\delta)},(\rho)}(V^r) \right).$$

Checking easily that  $\frac{\Gamma_n^{(1+\delta)}}{\Gamma_n^{1+\delta}} \xrightarrow{n \rightarrow +\infty} 0$ , (4.28) follows from Proposition 4.4(ii).  $\square$

**LEMMA 4.3.** *Let  $a \in (0, 1]$  such that  $(\mathbf{S}_a)$  holds. Assume that  $(X_t)$  admits a unique invariant distribution  $\nu$ . Assume  $(\mathbf{C}(\mathbf{f}, \mathbf{k}))$  and that  $\Gamma_n^{(3)} \xrightarrow{n \rightarrow +\infty} +\infty$ . Then,*

(i) *If  $\varphi_1$  defined by (4.10) satisfies  $(\mathbf{C}(\varphi_1, \mathbf{5}))$  then,*

$$\frac{1}{\Gamma_n^{(3)}} \mathcal{E}_{n,1} \xrightarrow{\rightarrow +\infty} -\frac{1}{2} m_{g_{\varphi_1}}^{(1)} \quad a.s.$$

(ii) *If the derivatives of  $g$  up to order 6 are continuous and dominated by  $V^r$  (with  $r > 0$ ),*

$$\frac{1}{\Gamma_n^{(3)}} \mathcal{E}_{n,2} \xrightarrow{\rightarrow +\infty} -\frac{1}{2} \nu(\varphi_2) \quad a.s.$$

*Proof.* (i) Writing

$$2 \sum_{k=1}^n \left( \frac{\gamma_k}{2} \right)^2 (\varphi_1(\bar{Y}_{2(k-1)}) + \mathbb{E}[\varphi_1(\bar{Y}_{2k-1}) | \mathcal{F}_{k-1}]) = \sum_{k=1}^n \frac{\gamma_k^2}{2} (\varphi_1(\bar{Y}_{2(k-1)}) + \varphi_1(\bar{Y}_{2k-1})) + \sum_{k=1}^n \frac{\gamma_k^2}{2} \Delta T_k$$

with  $\Delta T_k$  being a martingale increment defined by  $\Delta T_k = \mathbb{E}[\varphi_1(\bar{Y}_{2k-1}) | \mathcal{F}_{k-1}] - \varphi_1(\bar{Y}_{2k-1})$ , one obtains that

$$\mathcal{E}_{n,1} = \Gamma_n^{(2)} \left[ (\nu_n^{\gamma^{(2)},(\rho)} - \nu)(\varphi_1) - (\nu_n^{\gamma^{(2)}} - \nu)(\varphi_1) \right] + \sum_{k=1}^n \frac{\gamma_k^2}{2} \Delta T_k.$$

Applying Theorem V.3 of [Lem05] (which is an extension of Proposition 4.4(iv) to general weights) with  $\eta_k = \gamma_k^2$  and  $q^* = 4$ , we obtain that

$$\frac{\Gamma_n^{(2)}}{\Gamma_n^{(3)}} (\nu_n^{\gamma^{(2)},(\rho)} - \nu)(\varphi_1) \xrightarrow{n \rightarrow +\infty} m_{g_{\varphi_1}} \in \mathbb{R} \quad \text{in probability.}$$

Similarly, applying this result to the Euler scheme with half-step, we have:

$$\frac{\Gamma_n^{(2)}}{\Gamma_n^{(3)}} [(\nu_n^{\gamma^{(2)},(\rho)} - \nu)(\varphi_1)] = \frac{1}{2} \frac{\Gamma_n^{(2)}}{\sum_{k=1}^n \gamma_k^2 \cdot \frac{\gamma_k}{2}} [(\nu_n^{\gamma^{(2)},(\rho)} - \nu)(\varphi_1)] \xrightarrow{n \rightarrow +\infty} \frac{1}{2} m_{g_{\varphi_1}} \in \mathbb{R} \quad \text{in probability.}$$

Thus, it remains to show that the martingale term is negligible. We set  $\theta_k = \frac{\gamma_k^3}{\Gamma_k^{(3)^2}}$ . Using that  $(\gamma_k)$  is non-increasing, one checks that  $(\theta_n)$  is non-increasing and that  $\sum \theta_k \gamma_k < +\infty$ . Since  $|\varphi_1| \leq CV^r$  with  $r > 0$ , it follows from Proposition 4.4 that

$$\sum_{k \geq 1} \frac{\gamma_k^4}{(\Gamma_k^{(3)})^2} \mathbb{E}[|\varphi_1|^2(\bar{Y}_{2k-1})] < +\infty.$$

This implies that the martingale  $\sum \frac{\gamma_k^2}{\Gamma_k^{(3)}} \Delta T_k$  is *a.s.* convergent so that the Kronecker lemma yields  $\frac{1}{\Gamma_n^{(3)}} \sum_{k=1}^n \gamma_k^2 \Delta T_k \xrightarrow{n \rightarrow +\infty} 0$  *a.s.*. The first assertion follows.

(ii) Remark that

$$\mathcal{E}_{n,2} = \frac{1}{2} \nu_n^{\gamma^{(3)},(\rho)} - \nu_n^{\gamma^{(3)}}(\omega, \varphi_2) + \sum_{k=1}^n \frac{\gamma_k^3}{4} T_k.$$

Under the assumptions,  $\varphi_2$  is continuous and dominated by  $V^r$  with a positive  $r$ . Then, since  $\Gamma_n^{(3)} \xrightarrow{n \rightarrow +\infty} +\infty$ ,  $(\nu_n^{\gamma^{(3)},(\rho)}(\varphi_2))_{n \geq 1}$  and  $(\nu_n^{\gamma^{(3)}}(\omega, \varphi_2))_{n \geq 1}$  converge to  $\nu(\varphi_2)$ . With some similar arguments as previously, one checks that the martingale term is negligible and the second assertion follows.  $\square$

### 4.3.2 Proof of Theorem 4.3

For the sake of simplicity, we choose to give the proof of Theorem 4.3 only when  $\Gamma_n^{(3)} \xrightarrow{n \rightarrow +\infty} +\infty$ . Note that if  $\gamma_n = Cn^{-\mu}$ , this corresponds to  $\mu \leq 1/3$ , *i.e.* the case where the Romberg extrapolation really increases the rate of convergence (see Remark 4.8).

By the decomposition of Lemma 4.1 and the convergences established in Lemmas 4.2 and 4.3, one checks that it is now enough to prove the following points:

$$\frac{\Theta_n}{\Gamma_n} (2(g(\bar{Y}_{2n}) - g(\bar{Y}_0)) - (g(\bar{X}_n) - \bar{X}_0)) \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow +\infty, \quad (4.29)$$

$$\frac{\Theta_n}{\Gamma_n} N_n \xrightarrow{\mathbb{P}} 0 \quad \text{and} \quad \frac{\Theta_n}{\Gamma_n} R_n \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow +\infty \quad (4.30)$$

with  $\Theta_n = \sqrt{\Gamma_n} \vee \frac{\Gamma_n}{\Gamma_n^{(3)}}$ .

For (4.29), the result is obvious when  $g$  is bounded. Otherwise, we use Lemma 3 of [LP03] which implies in particular that for every  $p > 0$ ,  $\mathbb{E}[V^p(\bar{X}_n)] \leq C_p \Gamma_n$ . By Jensen's inequality, this implies that for every  $r > 0$  and  $\alpha \in (0, 1]$ , there exists a constant  $C > 0$  such that

$$\forall n \geq 1, \quad \mathbb{E}[V^r(\bar{X}_n)] \leq (\mathbb{E}[V^{\frac{r}{\alpha}}(\bar{X}_n)])^\alpha \leq C \frac{\Gamma_n}{\alpha} \Gamma_n^\alpha.$$

Thus, since the same property holds for the  $(\bar{Y}_n)$  and since  $|g| \leq CV^r$  with  $r > 0$ , (4.30) follows taking  $\alpha \in (0, 1/2)$ .

For the first assertion of (4.30), we use a martingale argument. We denote by  $\{\pi_{k,n}, k = 1, \dots, n, n \geq 1\}$  the triangular array of  $(\mathcal{F}_k)$ -martingale increments defined by

$$\pi_{k,n} = \frac{\Delta N_k}{\sqrt{\Gamma_n}}.$$

Then, in order to prove the convergence in probability of  $(N_n/\sqrt{\Gamma_n})$  to 0, we use the CLT for martingale increments which says that, since a Lindeberg-type condition holds (we do not prove this point, see Proof of Lemma 4.2 for a similar argument), it is enough to show that

$$\sum_{k=1}^n \mathbb{E}[|\pi_{k,n}|^2 | \mathcal{F}_{k-1}] \xrightarrow{n \rightarrow +\infty} 0 \quad \text{a.s.} \quad (4.31)$$

Under the assumptions on  $g$  and on the coefficients, one checks that there exists  $r > 0$  such that

$$\sum_{k=1}^n \mathbb{E}[|\pi_{k,n}|^2 | \mathcal{F}_{k-1}] \leq C \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k^2 (V^r(\bar{X}_{k-1}) + V^r(\bar{Y}_{2(k-1)}) + V^r(\bar{Y}_{2k-1})).$$

By Proposition 4.4,  $\sup_{n \geq 1} (\nu_n(\omega, V^r) + \nu_n^{(\rho)}(\omega, V^r)) < +\infty$ . Assertion (4.31) follows.

As concerns  $R_n$ , it follows from a martingale argument that

$$\frac{1}{\Gamma_n^{(3)}} \sum_{k=1}^n (\Delta R_k - \mathbb{E}[\Delta R_k | \mathcal{F}_{k-1}]) \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow +\infty.$$

Now, since  $\sup_{n \geq 1} (\nu_n^{\gamma^3}(\omega, V^r) + \nu_n^{\gamma^3, (\rho)}(\omega, V^r)) < +\infty$  *a.s.* and since  $\gamma_n \xrightarrow{n \rightarrow +\infty} 0$ , we deduce that

$$\frac{1}{\Gamma_n^{(3)}} \sum_{k=1}^n \mathbb{E}[\Delta R_k | \mathcal{F}_{k-1}] \xrightarrow{\mathbb{P}} 0.$$

The last assertion follows.

## A Hypo-ellipticity of the correlated duplicated system

It is a well-known fact that, for a Markov process, the strong Feller property combined with some irreducibility of the transitions implies uniqueness of the invariant distribution (see *e.g.* [DPZ96], Theorem 4.2.1). For a diffusion process with smooth coefficients, such properties hold if it satisfies the hypoelliptic Hörmander assumption (see [Hör67, Hör85]) and if the deterministic system related to the stochastic differential system (written in the Stratonovich sense) is controllable. In fact, both properties can be transferred from the original *SDE* to the duplicated system so that its invariant distribution is also unique. The main result of this section is Proposition A.6. Before, we need to introduce some Hörmander-type notations. First, written in a Stratonovich way,  $X$  is a solution to

$$dX_t = A_0(X_t)dt + \sum_{j=1}^q A_j(X_t) \circ dW_t^j \quad (\text{A.32})$$

where  $A_0, \dots, A_q$  denote some vectors fields on  $\mathbb{R}^d$  defined by<sup>1</sup>:

$$A_0(x) = \sum_{i=1}^d \left[ b_i(x) - \frac{1}{2} \sum_{l,j} \sigma_{l,j}(x) \partial_{x_j} \sigma_{i,l}(x) \right] \partial_{x_i}$$

and for every  $j \in \{1, \dots, q\}$ :

$$A_j(x) = \sum_{i=1}^d \sigma_{i,j}(x) \partial_{x_i}.$$

---

<sup>1</sup>With a standard abuse of notation, we identify the vectors fields and the associated differential operators.

For the sake of simplicity, we assume that  $b$  and  $\sigma$  are  $C^\infty$  on  $\mathbb{R}^d$  with bounded derivatives. We will also assume the following Hörmander condition at each point: there exists  $N \in \mathbb{N}^*$  such that  $\forall x \in \mathbb{R}^d$ ,

$$\dim(\text{Span}\{A_1(x), A_2(x), \dots, A_q(x), \text{L. B. of length } \leq N \text{ of the } A_j(x)\text{'s}, 0 \leq j \leq q\}) = d \quad (\text{A.33})$$

where ‘‘L.B.’’ stands for Lie Brackets. The above assumptions imply that for every  $t > 0$  and  $x \in \mathbb{R}^d$ ,  $P_t(x, \cdot)$  admits a density  $p_t(x, \cdot)$  w.r.t. the Lebesgue measure and that  $(x, y) \mapsto p_t(x, y)$  is  $C^\infty$  on  $\mathbb{R}^d \times \mathbb{R}^d$  (see e.g. [Cat92], Theorem 2.9). In particular,  $x \mapsto P_t(x, \cdot)$  is a strong Feller semi-group. Assume also that the control system (associated with (A.32))

$$\dot{x}^{(u)} = A_0(x^{(u)}) + \sum_{j=1}^q A_j(x^{(u)})u_j, \quad (\text{A.34})$$

is *approximately-controllable*:

There exists  $T > 0$  such that for every  $\varepsilon > 0$ ,  $x_1, x_2 \in \mathbb{R}^d$ , there exists  $u \in L^2([0, T], \mathbb{R}^d)$  such that  $(x^{(u)}(t))$  solution to (A.34) satisfies  $x(0) = x_1$  and  $|x(T) - x_2| \leq \varepsilon$ . (A.35)

Under Assumptions (A.33) and (A.35), the diffusion has a unique invariant distribution  $\nu$ . Actually, the controllability assumption combined with the Support Theorem implies that for every non-empty open set  $O$ , for every  $x \in \mathbb{R}^d$ ,  $P_T(x, O) > 0$ . The semi-group  $(P_t)$  is then irreducible. Owing to the strong Feller property, it follows classically that  $(P_t)$  admits an unique invariant distribution (see e.g. [DPZ96], Proposition 4.1.1. and Theorem 4.2.1.).

Furthermore,  $\nu$  is absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}^d$  and its topological support is  $\mathbb{R}^d$  (since for every open set  $O$  of  $\mathbb{R}^d$ ,  $\nu(O) = \int P_T(x, 0)\nu(dx) > 0$ ). Let us now consider the duplicated diffusion  $(X_t, X_t^{(\rho)})$ . Setting  $Z_t^{(\rho)} = (X_t, X_t^{(\rho)})$  and using the preceding notations, (1.2) can be written:

$$dZ_t^{(\rho)} = \tilde{A}_0(Z_t^{(\rho)})dt + \sum_{j=1}^q \tilde{A}_j(Z_t^{(\rho)})dW_t^j + \sum_{j=1}^q \tilde{A}_{d+j}(Z_t^{(\rho)})d\tilde{W}_t^j$$

where  $\tilde{A}_0(z) = (A_0(x), A_0(y))^T$  (with  $A_0(y) = \sum_{i=1}^d \left[ b_i(y) - \frac{1}{2} \sum_{l,j} \sigma_{l,j}(y) \partial_{y_j} \sigma_{i,l}(y) \right] \partial_{y_i}$  and  $z = (x, y)$ ),  $\tilde{W}$  is a  $d$ -dimensional Brownian Motion independent of  $W$  such that  $W^{(\rho)} = \rho^* W + (I_q - \rho^* \rho)^{\frac{1}{2}} \tilde{W}$  and for every  $j \in \{1, \dots, q\}$ ,

$$\tilde{A}_j(z) = A_j(x) + A_j^{(\rho)}(y) \quad \text{and} \quad \tilde{A}_{q+j}(z) = A_j^{((I_q - \rho^* \rho)^{\frac{1}{2}})}(y) \quad (\text{A.36})$$

where for a  $q \times q$  matrix  $B$ ,  $A_j^{(B)}(y) = \sum_{i=1}^d (\sigma(y)B)_{i,j} \partial_{y_i}$ . Then, the following property holds.

**PROPOSITION A.6.** *Let  $\rho \in \mathbb{M}_{q,q}(\mathbb{R})$  such that  $\rho^* \rho < I_q$ . Assume that  $b$  and  $\sigma$  are  $C^\infty$  on  $\mathbb{R}^d$  with bounded derivatives. Assume (A.33) and (A.35). Then, uniqueness holds for the invariant distribution  $\nu^{(\rho)}$  of the duplicated diffusion  $(X_t, X_t^{(\rho)})$ . Furthermore, if  $\nu^{(\rho)}$  exists, then  $\nu^{(\rho)}$  has a density  $p^{(\rho)}$  (w.r.t.  $\lambda_{2d}$ ) which is a.s. positive.*



*Proof.* First, let us check the Hörmander conditions for  $(X_t, X_t^{(\rho)})_{t \geq 0}$ . Setting  $S = (I_q - \rho^* \rho)^{\frac{1}{2}}$ , standard computations yield

$$\forall j \in \{1, \dots, q\}, \quad \tilde{A}_{q+j}(z) = \sum_{l=1}^q S_{l,j} A_l(y).$$

Since  $S$  is invertible, we deduce that  $\{A_l(y), l = 1, \dots, q\}$  belongs to  $\text{Span}\{\tilde{A}_{q+j}(z), j = 1, \dots, q\}$ . Similarly, checking that for every  $j \in \{1, \dots, q\}$ ,

$$[\tilde{A}_0(z), \tilde{A}_{q+j}(z)] = [A_0(y), A_j^{(S)}(y)] = \sum_{l=1}^q S_{l,j} [A_0(y), A_l(y)]$$

one deduces from the invertibility of  $S$  that  $\{[A_0(y), A_l(y)], l = 1, \dots, q\}$  is included in  $\text{Span}\{[\tilde{A}_0(z), \tilde{A}_{q+j}(z)], j = 1, \dots, q\}$ . Owing to (A.33), it follows that  $\text{Span}\{\partial_{y_1}, \dots, \partial_{y_d}\}$  is included in

$$V = \text{Span} \left\{ \tilde{A}_1(z), \tilde{A}_2(z), \dots, \tilde{A}_q(z), \text{Lie Brackets of length} \leq N \text{ of the } \tilde{A}_j(z) \text{'s}, 0 \leq j \leq q \right\}.$$

Now, let us show that  $\text{Span}\{\partial_{x_1}, \dots, \partial_{x_d}\}$  is included in  $V$ . Since  $\text{Span}\{\partial_{y_1}, \dots, \partial_{y_d}\}$  is included in  $V$ , it is clear that for every  $x \in \mathbb{R}^d$ ,  $A_j(x) = A_j^{(\rho)}(y) - \tilde{A}_j(z)$  also belongs to  $V$ . Since

$$[\tilde{A}_0(z), \tilde{A}_j(z)] = [A_0(x), A_j(x)] + [A_0(y), A_j^{(\rho)}(y)],$$

$[A_0(x), A_j(x)]$  has the same property. Using again (A.33), we deduce that  $\text{Span}\{\partial_{x_1}, \dots, \partial_{x_d}\}$  is included in  $V$  and thus that  $\dim(V) = 2d$ . As a consequence, for every  $z \in \mathbb{R}^{2d}$  and  $t > 0$ ,  $Q_t^{(\rho)}(z, \cdot)$  admits a density  $q_t(z, \cdot)$  w.r.t.  $\lambda_{2d}$  such that  $(z, z') \mapsto q_t(z, z')$  is  $\mathcal{C}^\infty$  on  $\mathbb{R}^{2d} \times \mathbb{R}^{2d}$ .

In order to obtain uniqueness for the invariant distribution, it remains to show that there exists  $T > 0$  such that for every  $z \in \mathbb{R}^{2d}$ , for every non-empty open set  $O$  of  $\mathbb{R}^{2d}$ ,  $Q_T(z, O) > 0$ . Owing to (A.35), it is clear that for every  $z_1 = (x_1, y_1)$  and  $z_2 = (x_2, y_2)$ , for every  $\varepsilon > 0$ , there exist  $u$  and  $\tilde{u} \in L^2([0, T], \mathbb{R}^d)$  such that  $z(t) = (x^{(u)}(t), x^{(\tilde{u})}(t))$ , where  $x^{(u)}$  and  $x^{(\tilde{u})}$  are solutions to (A.34) starting from  $x_1$  and  $y_1$ , satisfies  $|z(T) - z_2| \leq \varepsilon$ . Furthermore, since  $S$  is invertible, we can assume that  $\tilde{u} = \rho u + S\omega$  with  $\omega \in L^2([0, T], \mathbb{R}^d)$ . Then, the support Theorem can be applied to obtain that for every  $z_1, z_2, \varepsilon$   $Q_T(z_1, B(z_2, \frac{\varepsilon}{2})) > 0$  and thus to conclude that for every  $z \in \mathbb{R}^{2d}$  and every non-empty open set  $O$ ,  $Q_T(z, O) > 0$ .  $\square$

## B Additional proofs about the two-dimensional counterexample

**Proof of (3.5):** For the sake of completeness, we show that  $r_t \rightarrow 1$  a.s. as soon as  $r_0 > 0$ . First, note that uniqueness holds for the solution of the SDE (3.3) since the coefficients are Lipschitz continuous. In particular,  $(r_t^1)$  defined a.s. by  $r_t^1 = 1$  for every  $t \geq 0$  is the unique solution starting from  $r_0 = 1$ . Owing to the strong Markov property, this implies that if  $\tau^1 := \inf\{t \geq 0, r_t = 1\}$ , then  $r_t = 1$  on  $\{\tau \leq t\}$ . The same property holds at 0. We deduce that  $(r_t)_{t \geq 0}$  lives in  $[1, +\infty)$  if  $r_0 > 1$  and in  $[0, 1]$  if  $r_0 \in [0, 1]$ . Moreover, if  $r_0 > 1$ , we have  $d(r_t - 1) = -(r_t - 1)(dt + \vartheta dW_t)$  so that

$$r_t - 1 = e^{-(1 + \frac{\vartheta^2}{2})t + \vartheta W_t}.$$

It follows that  $\lim_{t \rightarrow +\infty} r_t = 1$  since  $\lim_{t \rightarrow +\infty} \frac{W_t}{t} = 0$  *a.s.*. Now, if  $r_0 \in [0, 1]$ , we have

$$dr_t = r_t(1 - r_t)(dt + \vartheta dW_t).$$

Thus,  $(r_t)$  is a  $[0, 1]$ -valued submartingale. In particular,  $r_t$  converges *a.s.* to a  $[0, 1]$ -valued random variable  $r_\infty$ . Since

$$\forall t \geq 0, \quad \mathbb{E}[r_t] = r_0 + \mathbb{E}\left(\int_0^t r_s(1 - r_s)ds\right),$$

it follows that  $\mathbb{E}[\int_0^{+\infty} r_s(1 - r_s)ds]$  which in turn implies that  $\int_0^{+\infty} r_s(1 - r_s)ds < +\infty$  *a.s.* As a consequence  $\liminf_{t \rightarrow +\infty} r_t(1 - r_t) = 0$  *a.s.*. The process  $(r_t)$  being *a.s.* convergent to  $r_\infty$ , it follows that  $r_\infty \in \{0, 1\}$  *a.s.*. It remains to prove that  $\mathbb{P}(r_\infty = 0) = 0$ . Denote by  $p$  the scale function of  $(r_t)$  null at  $r = 1/2$ . For every  $r \in (0, 1)$ ,

$$p(r) = \int_{\frac{1}{2}}^r e^{-\int_{\frac{1}{2}}^\xi \frac{2}{\vartheta^2 u(1-u)} du} d\xi = \int_{\frac{1}{2}}^r \left(\frac{1-\xi}{\xi}\right)^{\frac{2}{\vartheta^2}} d\xi.$$

As a consequence, if  $\vartheta \in (0, \sqrt{2}]$ ,  $\lim_{r \rightarrow +\infty} p(r) = +\infty$ . This means that 0 is a repulsive point and that, as a consequence (see *e.g.* [KT81], Lemma 6.1 p. 228),

$$\forall b \in (0, 1) \quad \mathbb{P}\left(\lim_{a \rightarrow 0^+} \tau_a < \tau_b\right) := \lim_{a \rightarrow 0^+} \mathbb{P}(\tau_a < \tau_b) = 0$$

where  $\tau_a = \inf\{t \geq 0 \mid r_t = a\}$ ,  $y \in [0, 1]$ . We deduce that  $\mathbb{P}(r_\infty = 0) = 0$ . This completes the proof.  $\square$

**Proof of (3.6):** We want to prove that  $\mu$  is invariant for  $(X_t^x, X_t^{x'})$  if and only if  $\mu$  can be represented by (3.6). First, since the unique invariant distribution of  $(X_t^x)$  is  $\lambda_{S_1}$ , it is clear that  $\mu = \mathcal{L}(e^{i\Theta_0}, e^{i(\Theta_0 + V_0)})$  where  $\Theta_0$  has uniform distribution on  $[0, 2\pi]$  and  $V_0$  is a random variable with values in  $[0, 2\pi)$ . One can check that if  $V_0$  is independent of  $\Theta_0$ ,  $\mu$  is invariant. Thus, it remains to prove that it is a necessary condition or equivalently that  $K(\theta, dv) := \mathcal{L}(e^{iV_0} \mid e^{i\Theta_0} = e^{i\theta})$  does not depend on  $\theta$ . Denote by  $(e^{i\Theta_t}, e^{i(\Theta_t + V_t)})$  the (stationary) duplicated diffusion starting from  $(e^{i\Theta_0}, e^{i(\Theta_0 + V_0)})$ . Since  $\mu$  is invariant, we have for every  $t \geq 0$

$$\mathcal{L}(e^{iV_t} \mid e^{i\Theta_t} = e^{i\theta}) = K(\theta, dv)$$

but thanks to the construction, for every  $t \geq 0$ ,  $\Theta_t = \Theta_0 + W_t$  and  $V_t = V_0$  (the angular difference between the two coordinates does not change) so that

$$\mathcal{L}(e^{iV_t} \mid e^{i\Theta_t} = e^{i\theta}) = \int K(\theta', dv) \rho_t(\theta, d\theta')$$

where  $\rho_t(\theta, d\theta') = \mathcal{L}(e^{i(\theta + W_t)})$ . But  $\rho_t(\theta, d\theta')$  converges weakly to  $\lambda_{S_1}$  when  $t \rightarrow +\infty$ . From the two previous equations it follows that  $K(\theta, dv)$  does not depend on  $\theta$  since  $\forall \theta \geq 0$ ,  $K(\theta, dv) = \int K(\theta', dv) \lambda_{S_1}(d\theta')$ .  $\square$

## References

- [BB92] G. K. Basak and R. N. Bhattacharya. Stability in distribution for a class of singular diffusions. *Ann. Probab.*, 20(1):312–321, 1992.

- [Bha82] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrsch. Verw. Gebiete*, 60(2):185–201, 1982.
- [Cat92] P. Cattiaux. Stochastic calculus and degenerate boundary value problems. *Ann. Inst. Fourier (Grenoble)*, 42(3):541–624, 1992.
- [CL89] M. F. Chen and S. F. Li. Coupling methods for multidimensional diffusion processes. *Ann. Probab.*, 17(1):151–177, 1989.
- [DPZ96] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.
- [Har81] T. E. Harris. Brownian motions on the homeomorphisms of the plane. *Ann. Probab.*, 9(2):232–254, 1981.
- [Has80] R. Z. Has'minskiĭ. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.
- [HH80] P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. Probability and Mathematical Statistics.
- [Hör67] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
- [Hör85] L. Hörmander. *The analysis of linear partial differential operators 1-4.*, volume 275 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1985.
- [IW77] N. Ikeda and S. Watanabe. A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka J. Math.*, 14(3):619–633, 1977.
- [KT81] S. Karlin and H. M. Taylor. *A second course in stochastic processes*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1981.
- [Kun90] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.
- [Lem05] V. Lemaire. *Estimation récursive de la mesure invariante d'un processus de diffusion*. Thèse de doctorat, Université de Marne-la-Vallée (France), 2005.
- [LP02] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8(3):367–405, 2002.
- [LP03] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion: the case of a weakly mean reverting drift. *Stoch. Dyn.*, 3(4):435–451, 2003.
- [Pan06] F. Panloup. *Approximation récursive du régime stationnaire d'une Equation Différentielle Stochastique avec sauts*. Thèse de doctorat, Université Pierre et Marie Curie (Paris 6, France), 2006.
- [PV01] E. Pardoux and A. Yu. Veretennikov. On the Poisson equation and diffusion approximation. I. *Ann. Probab.*, 29(3):1061–1085, 2001.