



Rate of convergence to equilibrium for discrete-time stochastic dynamics with memory

Maylis Varvenne

► To cite this version:

Maylis Varvenne. Rate of convergence to equilibrium for discrete-time stochastic dynamics with memory. 2017. hal-01588505v1

HAL Id: hal-01588505

<https://hal.science/hal-01588505v1>

Preprint submitted on 15 Sep 2017 (v1), last revised 12 Nov 2018 (v4)

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.

Rate of convergence to equilibrium for discrete-time stochastic dynamics with memory

Maylis Varvenne*

September 15, 2017

Abstract

The main objective of the paper is to study the long-time behavior of general discrete dynamics driven by ergodic stationary Gaussian noise. To this end, we first explain how is it possible to define invariant distributions in this generally non-Markovian setting and to get existence results under appropriate conditions. Then, we get a uniqueness result and a rate of convergence to the invariant distribution in total variation thanks to a coupling procedure (with a step specific to non-Markovian framework).

Keywords: Discrete stochastic dynamics; Stationary Gaussian noise; Rate of convergence to equilibrium; Total variation distance; Lyapunov function; Toeplitz operator.

1 Introduction

Convergence to equilibrium for Stochastic dynamics is one of the most natural and most studied problems in probability theory. Regarding Markov processes, this topic has been deeply undertaken through various approaches: spectral analysis, functional inequalities or coupling methods. However, in many applications (Physics, Biology, Finance...) the future evolution of a quantity may depend on its own history, and thus, noise with independent increments does not accurately reflect reality. A classical way to overcome this problem is to consider dynamical systems driven by processes with stationary increments like fractional Brownian motion for instance. In a continuous time framework, Stochastic Differential Equations (SDEs) driven by Gaussian processes with stationary increments have been introduced to model random evolution phenomena with long range dependence properties. Consider SDEs of the following form

$$dX_t = b(X_t) + \sigma(X_t)dZ_t \quad (1.1)$$

where $(Z_t)_{t \geq 0}$ is a Gaussian process with ergodic stationary increments and $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_d(\mathbb{R})$, $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ are functions defined in a such a way that existence of a solution holds. As concerns long-time behavior, different properties have been studied like approximation of stationary solution in [8] or the rate of convergence to an equilibrium distribution. For this last property, the case when $(Z_t)_{t \geq 0}$ is fractional Brownian motion (fBm) has received significant attention from Hairer [14], Fontbona and Panloup [11], Deya, Panloup and Tindel [9] over the last decade. They used coalescent coupling strategy to compute the rate of convergence. In the additive noise setting, Hairer proved that the process converges in total variation to the stationary regime with a rate upper-bounded by $C_\varepsilon t^{-(\alpha_H - \varepsilon)}$ for any $\varepsilon > 0$, with

$$\alpha_H = \begin{cases} \frac{1}{8} & \text{if } H \in (\frac{1}{4}, 1) \setminus \{\frac{1}{2}\} \\ H(1 - 2H) & \text{if } H \in (0, \frac{1}{4}). \end{cases} \quad (1.2)$$

In the multiplicative noise setting, Fontbona and Panloup extended those results under selected assumptions on σ to the case where $H \in (\frac{1}{2}, 1)$ and finally Deya, Panloup and Tindel obtained this type of results

*Institut de Mathématiques de Toulouse, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France.
E-mail: maylis.varvenne@math.univ-toulouse.fr

in the rough setting $H \in (\frac{1}{3}, 1/2)$. Here, we focus on a general class of recursive discrete dynamics driven by a stationary sequence which includes in particular discretization of (1.1), that is

$$X_{n+1} = F(X_n, \Delta_{n+1}) \quad (1.3)$$

where $(\Delta_n)_{n \in \mathbb{Z}}$ is an ergodic stationary Gaussian sequence. This type of discrete stochastic dynamics, which is not Markovian in general, has not been much discussed except in the linear case like Autoregressive Moving-average (ARMA) models [5] whose main objective is the prediction of stationary processes. When F is linear, dynamics like (1.3) are truly related to ARMA processes through the so-called Wold's decomposition theorem which implies that we can see $(\Delta_n)_{n \in \mathbb{Z}}$ as a moving-average of infinite order (see [5] to get more details).

Here, we investigate the problem of the long-time behavior of (1.3) for a general class of functions F . To this end, we first explain how is it possible to define invariant distributions in this non-Markovian setting and to obtain existence results, and then we use a coalescent coupling strategy to get the rate of convergence to equilibrium of processes following (1.3) under selected assumptions. This discrete time framework has several advantages. First, it allows us to better target the impact of the memory thanks to the moving average (MA) representation of the noise process (see (2.2)). The deterministic sequence defined by the coefficients involved in this representation measures, in a sense, the weight of the past since the covariance function of $(\Delta_n)_{n \in \mathbb{Z}}$ is entirely determined by those coefficients (see Remark 2.1). Then, one of our motivations to work in this discrete context is to see if the speed of convergence to equilibrium is affected by the difficulty of the coupling strategy which is a priori greater in a continuous time setting.

Now, let us briefly recall how this coupling method is organized. First, one takes two processes $(X_n^1, (\Delta_{n+k}^1)_{k \leq 0})_{n \geq 0}$ and $(X_n^2, (\Delta_{n+k}^2)_{k \leq 0})_{n \geq 0}$ following (1.3) starting respectively from μ_0 and μ_* (the invariant distribution). As a preliminary step, one waits that the two paths get close. Then, at each trial, the coupling attempt is divided in two steps. First, one tries in Step 1 to stick the positions together at a given time. Then, in Step 2, one tries to ensure that the paths stay clustered until $+\infty$. Actually, oppositely to the Markovian setting where the paths remain naturally fastened together (by putting the same innovation on each path), the main difficulty here is that, staying together has a cost. In other words, this property can be ensured only with a non trivial coupling of the noises. Finally, if one of the two previous steps fails, one begins Step 3 by putting the same noise on each coordinate until the "cost" to attempt Step 1 is not too big. In other words, during this step one waits again for the paths to get close but also for the memory of the coupling cost to decrease sufficiently.

Thanks to this strategy, we are able to prove that the law of the process $(X_{n+k})_{k \geq 0}$ following (1.3) converges in total variation to the stationary regime with a rate upper-bounded by Cn^{-v} where v is a quantity which is directly linked to the asymptotic behavior of the sequence of coefficients appearing in the MA representation of the noise process. In particular, we focus on Gaussian noise with exponential and polynomial memory (see Section 2 for more details). For the polynomial case, a more precise example is also studied, namely noise with fractional memory (see Subsection 2.6). This example coupled with the fact that we apply our result to the discretization of (1.1) (see Subsection 2.4) allows us to contrast with the continuous time results [14, 11, 9].

The following section gives more details on the studied dynamics and describes the assumptions required to get the main result, namely Theorem 2.1. Then, the proof of Theorem 2.1 is achieved in Sections 3, 4, 5, 6 and 7, which are outlined at the end of Section 2.

2 Setting and main results

2.1 Setting

Let $X := (X_n)_{n \geq 0}$ denote an \mathbb{R}^d -valued sequence defined by: X_0 is a random variable with distribution denoted by μ_0 and

$$\forall n \geq 0, \quad X_{n+1} = F(X_n, \Delta_{n+1}), \quad (2.1)$$

where $F : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is continuous and $(\Delta_n)_{n \in \mathbb{Z}}$ is a stationary and purely non-deterministic Gaussian sequence. Hence, by Wold's decomposition theorem [5] it has a moving average representation

$$\forall n \in \mathbb{Z}, \quad \Delta_n = \sum_{k=0}^{+\infty} a_k \xi_{n-k} \quad (2.2)$$

with

$$\begin{cases} (a_k)_{k \geq 0} \in \mathbb{R}^{\mathbb{N}} \text{ such that } a_0 \neq 0 \text{ and } \sum_{k=0}^{+\infty} a_k^2 < +\infty \\ (\xi_k)_{k \in \mathbb{Z}} \text{ an i.i.d sequence such that } \xi_1 \sim \mathcal{N}(0, I_d). \end{cases} \quad (2.3)$$

Without loss of generality, we assume that $a_0 = 1$. Actually, if $a_0 \neq 1$, we can come back to this case by setting

$$\tilde{\Delta}_n = \sum_{k=0}^{+\infty} \tilde{a}_k \xi_{n-k}$$

with $\tilde{a}_k := \frac{a_k}{a_0}$.

Remark 2.1. *The asymptotic behavior of the sequence $(a_k)_{k \geq 0}$ certainly plays a key role to compute the rate of convergence to equilibrium of the process $(X_n)_{n \geq 0}$. Actually, the memory induced by the noise process is quantified by the sequence $(a_k)_{k \geq 0}$ through the identity*

$$\forall n \in \mathbb{Z}, \forall k \geq 0, \quad c(k) := \mathbb{E}[\Delta_n \Delta_{n+k}] = \sum_{i=0}^{+\infty} a_i a_{k+i}.$$

In the sequel, the *state space* of the process X and the *noise space* associated to $((\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ will be respectively denoted by $\mathcal{X} := \mathbb{R}^d$ and $\mathcal{W} := (\mathbb{R}^d)^{\mathbb{Z}^-}$. These notations will be clarified in Subsection 3.1.

2.2 Preliminary tool: a Toeplitz type operator

The moving-average representation of the Gaussian sequence $(\Delta_n)_{n \in \mathbb{Z}}$ naturally leads us to define an operator related to the coefficients $(a_k)_{k \geq 0}$. First, set

$$\ell_a(\mathbb{Z}^-, \mathbb{R}^d) := \left\{ w \in (\mathbb{R}^d)^{\mathbb{Z}^-} \mid \forall k \geq 0, \sum_{l=0}^{+\infty} a_l w_{-k-l} < +\infty \right\}$$

and define \mathbf{T}_a on $\ell_a(\mathbb{Z}^-, \mathbb{R}^d)$ by

$$\mathbf{T}_a(w) = \left(\sum_{l=0}^{+\infty} a_l w_{-k-l} \right)_{k \geq 0}. \quad (2.4)$$

Due to the Cauchy-Schwarz inequality, we can check that for instance $\ell^2(\mathbb{Z}^-, \mathbb{R}^d)$ is included in $\ell_a(\mathbb{Z}^-, \mathbb{R}^d)$ due to the assumption $\sum_{k \geq 0} a_k^2 < +\infty$. This Toeplitz type operator \mathbf{T}_a links $(\Delta_n)_{n \in \mathbb{Z}}$ to $(\xi_n)_{n \in \mathbb{Z}}$. The following proposition spells out the reverse operator.

Proposition 2.1. *Let \mathbf{T}_b be the operator defined on $\ell_b(\mathbb{Z}^-, \mathbb{R}^d)$ in the same way as \mathbf{T}_a but with the following sequence $(b_k)_{k \geq 0}$*

$$b_0 = \frac{1}{a_0} \quad \text{and} \quad \forall k \geq 1, \quad b_k = -\frac{1}{a_0} \sum_{l=1}^k a_l b_{k-l}. \quad (2.5)$$

Then,

$$\forall w \in \ell_a(\mathbb{Z}^-, \mathbb{R}^d), \quad \mathbf{T}_b(\mathbf{T}_a(w)) = w \quad \text{and} \quad \forall w \in \ell_b(\mathbb{Z}^-, \mathbb{R}^d), \quad \mathbf{T}_a(\mathbf{T}_b(w)) = w$$

that is $\mathbf{T}_b = \mathbf{T}_a^{-1}$ and $\ell_b(\mathbb{Z}^-, \mathbb{R}^d) = \mathbf{T}_a(\ell_a(\mathbb{Z}^-, \mathbb{R}^d))$.

Proof. Let $w \in \ell_a(\mathbb{Z}^-, \mathbb{R}^d)$. Then let $n \geq 0$,

$$\begin{aligned}
(\mathbf{T}_b(\mathbf{T}_a(w)))_{-n} &= \sum_{k=0}^{+\infty} b_k (\mathbf{T}_a(w))_{-n-k} \\
&= \sum_{k=0}^{+\infty} b_k \sum_{l=0}^{+\infty} a_l w_{-n-k-l} \\
&= \sum_{k=0}^{+\infty} \sum_{i=k}^{+\infty} b_k a_{i-k} w_{-n-i} \quad (\text{by setting } i = k + l) \\
(\mathbf{T}_b(\mathbf{T}_a(w)))_{-n} &= \sum_{i=0}^{+\infty} \underbrace{\left(\sum_{k=0}^i b_k a_{i-k} \right)}_{=0 \text{ except for } i=0} w_{-n-i} = w_{-n}
\end{aligned}$$

We show in the same way that for $w \in \ell_b(\mathbb{Z}^-, \mathbb{R}^d)$, we have $\mathbf{T}_a(\mathbf{T}_b(w)) = w$. \square

Remark 2.2. The sequence $(b_k)_{k \geq 0}$ is of first importance in the sequel. The sketch of the proof of Theorem 2.1 will use an important property linked to the sequence $(b_k)_{k \geq 0}$: if two sequences u and v are such that

$$\forall n \geq 1, \quad u_n = \sum_{k=0}^{n-1} a_k v_{n-k}$$

then

$$\forall n \geq 1, \quad v_n = \sum_{k=0}^{n-1} b_k u_{n-k}.$$

This reverse identity and the asymptotic behavior of $(b_k)_{k \geq 0}$ play a significant role in the computation of the rate of convergence.

The following section is devoted to outline assumptions on $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ and then on F to get the main result.

2.3 Assumptions and general theorem

First of all, let us introduce assumptions on $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$. All along the paper, we will switch between two types of assumptions called respectively the *polynomial case* and the *exponential case*.

Hypothesis (\mathbf{H}_{poly}): The following conditions hold,

- there exist $\rho, \beta > 0$ and $C_\rho, C_\beta > 0$ such that

$$\forall k \geq 0, \quad |a_k| \leq C_\rho (k+1)^{-\rho} \quad \text{and} \quad \forall k \geq 0, \quad |b_k| \leq C_\beta (k+1)^{-\beta}.$$

- there exist $\kappa \geq \rho + 1$ and $C_\kappa > 0$ such that

$$\forall k \geq 0, \quad |a_k - a_{k+1}| \leq C_\kappa (k+1)^{-\kappa}.$$

Hypothesis (\mathbf{H}_{exp}): There exist $\lambda, \mu > 0$ and $C_\lambda, C_\mu > 0$ such that,

$$\forall k \geq 0, \quad |a_k| \leq C_\lambda e^{-\lambda k} \quad \text{and} \quad \forall k \geq 0, \quad |b_k| \leq C_\mu e^{-\mu k}.$$

Remark 2.3. \triangleright (\mathbf{H}_{poly}) and (\mathbf{H}_{exp}) are general parametric hypothesis which apply to a large class of Gaussian driven dynamics. These assumptions involve the memory of the noise process through the sequence $(a_k)_{k \geq 0}$ but also through the coefficients appearing in the reverse Toeplitz operator \mathbf{T}_a^{-1} (see

Proposition 2.1. Due to the strategy of the proof (coalescent coupling in a non Markovian setting) we also need a bound on the discrete derivative of $(a_k)_{k \geq 0}$.

▷ Even though $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ are related by (2.5), there is no general rule which connects ρ and β . This fact will be highlighted in Subsection 2.6.

Let us now introduce some assumptions on the function F which defines the dynamics (2.1). Throughout this paper $F : \mathcal{X} \times \mathbb{R}^d \rightarrow \mathcal{X}$ is a continuous function and the following hypothesis (\mathbf{H}_1) and (\mathbf{H}_2) are satisfied.

Hypothesis (\mathbf{H}_1) : There exists a continuous function $V : \mathcal{X} \rightarrow \mathbb{R}_+^*$ satisfying $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ and $\exists \gamma \in (0, 1)$ and $C > 0$ such that for all $(x, w) \in \mathcal{X} \times \mathbb{R}^d$,

$$V(F(x, w)) \leq \gamma V(x) + C(1 + |w|).$$

Remark 2.4. We will see in Subsection 3.2 that this condition on F ensures the existence of an invariant distribution (in a sense made precise below).

We define $\tilde{F} : \mathcal{X} \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathcal{X}$ by $\tilde{F}(x, u, y) = F(x, u + y)$. We assume that \tilde{F} satisfies the following conditions:

Hypothesis (\mathbf{H}_2) : Let $K > 0$. We assume that there exists $\tilde{K} > 0$ such that for every $\mathbf{x} := (x, x', y, y')$ in $B(0, K)^4$, there exist $\Lambda_{\mathbf{x}} : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $M_K > 0$ and $C_{\tilde{K}}$ such that the following holds

- $\Lambda_{\mathbf{x}}$ is a measurable, invertible and almost everywhere differentiable function such that $\Lambda_{\mathbf{x}}^{-1}$ is also measurable.
- for all $u \in B(0, \tilde{K})$,

$$\tilde{F}(x, u, y) = \tilde{F}(x', \Lambda_{\mathbf{x}}(u), y') \quad (2.6)$$

$$|\det(J_{\Lambda_{\mathbf{x}}}(u))| \geq C_{\tilde{K}} \quad (2.7)$$

- for all $u \in \mathbb{R}^d$,

$$|\Lambda_{\mathbf{x}}(u) - u| \leq M_K \quad (2.8)$$

Remark 2.5. Let us make a few precisions on the arguments of \tilde{F} : x is the position of the process, u the increment of the innovation process and y is related to the past of the process (see (4.7) for more details). The boundary $C_{\tilde{K}}$ and M_K are independent from x, x', y and y' . This assumption can be viewed as a kind of controlability assumption in the following sense: the existence of $\Lambda_{\mathbf{x}}$ leads to the coalescence of the positions by (2.6). This is of first importance to achieve the first step of the coupling procedure (see Subsection 4.2).

We are now in position to state our main result. Let $\mathcal{L}((X_n^{\mu_0})_{n \geq 0})$ denote the distribution of the process X starting from an initial condition μ_0 (see Subsection 3.1 below for detailed definitions of initial condition and invariant distribution) and for an invariant distribution μ_* denote by $S\mu_*$ the law of the stationary solution. Finally, we denote by $\|\cdot\|_{TV}$ the classical total variation norm.

Theorem 2.1. Assume (\mathbf{H}_1) and (\mathbf{H}_2) . Then,

- There exists an invariant distribution μ_* associated to (2.1).
- Assume that $(\mathbf{H}_{\text{poly}})$ is true with $\rho, \beta > 1/2$ and $\rho + \beta > 3/2$. Then, uniqueness holds for the invariant distribution μ_* . Furthermore, for every initial distribution μ_0 for which $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_0(dx) < +\infty$ and for all $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that

$$\|\mathcal{L}((X_{n+k}^{\mu_0})_{k \geq 0}) - S\mu_*\|_{TV} \leq C_\varepsilon n^{-(v(\beta, \rho) - \varepsilon)}.$$

where the function v is defined by

$$v(\beta, \rho) = \sup_{\alpha \in (\frac{1}{2} \vee (\frac{3}{2} - \beta), \rho)} \min\{1, 2(\rho - \alpha)\} (\min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2).$$

(iii) Assume that $(\mathbf{H}_{\text{exp}})$ is true, then uniqueness holds for the invariant distribution μ_\star . Furthermore, for every initial distribution μ_0 for which $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_0(dx) < +\infty$ and for all $p > 0$, there exists $C_p > 0$ such that

$$\|\mathcal{L}((X_{n+k}^{\mu_0})_{k \geq 0}) - \mathcal{S}\mu_\star\|_{TV} \leq C_p n^{-p}.$$

Remark 2.6. Assumption (\mathbf{H}_2) is only required to perform the first step of the coupling strategy to get (ii) and (iii) (see Section 4 for more details).

In the following subsection, we test the assumptions of our main result Theorem 2.1 (especially (\mathbf{H}_1) and (\mathbf{H}_2)) on the Euler scheme of SDEs like (1.1).

2.4 The Euler Scheme

Recall that $\mathcal{X} = \mathbb{R}^d$. In this subsection, set

$$\begin{aligned} F_h : \quad \mathcal{X} \times \mathbb{R}^d &\rightarrow \mathcal{X} \\ (x, w) &\mapsto x + hb(x) + \sigma(x)w. \end{aligned} \quad (2.9)$$

where $h > 0$, $b : \mathcal{X} \rightarrow \mathcal{X}$ is continuous and $\sigma : \mathcal{X} \rightarrow \mathcal{M}_d(\mathbb{R})$ is a continuous and bounded function on \mathcal{X} . For all $x \in \mathcal{X}$ we suppose that $\sigma(x)$ is invertible and we denote by $\sigma^{-1}(x)$ the inverse. Moreover, we assume that σ^{-1} is a continuous function and that b satisfies a Lyapunov type assumption that is:

(L1) $\exists C > 0$ such that

$$\forall x \in \mathcal{X}, \quad |b(x)| \leq C(1 + |x|) \quad (2.10)$$

(L2) $\exists \tilde{\beta} \in \mathbb{R}$ and $\tilde{\alpha} > 0$ such that

$$\forall x \in \mathcal{X}, \quad \langle x, b(x) \rangle \leq \tilde{\beta} - \tilde{\alpha}|x|^2. \quad (2.11)$$

Remark 2.7. This function F_h corresponds to the Euler scheme associated to SDEs like (1.1). The conditions on the function b are classical to get existence of invariant distribution.

In this setting the function \tilde{F}_h (introduced in Hypothesis (\mathbf{H}_2)) is given by

$$\begin{aligned} \tilde{F}_h : \quad \mathcal{X} \times \mathbb{R}^d \times \mathbb{R}^d &\rightarrow \mathcal{X} \\ (x, u, y) &\mapsto x + hb(x) + \sigma(x)(u + y). \end{aligned}$$

Theorem 2.2. Let $h > 0$. Let F_h be the function defined above. Assume that $b : \mathcal{X} \rightarrow \mathcal{X}$ is a continuous function satisfying **(L1)** and **(L2)** and $\sigma : \mathcal{X} \rightarrow \mathcal{M}_d(\mathbb{R})$ is a continuous and bounded function such that for all $x \in \mathcal{X}$, $\sigma(x)$ is invertible and $x \mapsto \sigma^{-1}(x)$ is a continuous function. Then, (\mathbf{H}_1) and (\mathbf{H}_2) hold for F_h with $h > 0$ small enough.

Proof. Set $V(x) = |x|$. Let us begin by proving that (\mathbf{H}_1) holds with V for F_h with $h > 0$ small enough. We have:

$$|F_h(x, w)|^2 = |x|^2 + h^2|b(x)|^2 + 2h\langle x, b(x) \rangle + 2\langle x, \sigma(x)w \rangle + 2h\langle b(x), \sigma(x)w \rangle + |\sigma(x)w|^2.$$

Then, using the inequality $|\langle a, b \rangle| \leq \frac{1}{2}(\varepsilon|a|^2 + \frac{1}{\varepsilon}|b|^2)$ for all $\varepsilon > 0$, we get

$$|\langle x, \sigma(x)w \rangle| \leq \frac{1}{2}(\varepsilon|x|^2 + \frac{1}{\varepsilon}|\sigma(x)w|^2) \quad \text{et} \quad |\langle b(x), \sigma(x)w \rangle| \leq \frac{1}{2}(\varepsilon|b(x)|^2 + \frac{1}{\varepsilon}|\sigma(x)w|^2).$$

Moreover, assumptions **(L1)** and **(L2)** on b give

$$|\langle b(x), x \rangle| \leq \tilde{\beta} - \tilde{\alpha}|x|^2 \quad \text{et} \quad |b(x)|^2 \leq \tilde{C}(1 + |x|^2).$$

Hence, we finally have

$$\begin{aligned}
|F_h(x, w)|^2 &\leq |x|^2 + \tilde{C}h^2(1 + |x|^2) + 2h(\tilde{\beta} - \tilde{\alpha}|x|^2) + \varepsilon|x|^2 + \frac{1}{\varepsilon}|\sigma(x)w|^2 \\
&\quad + \tilde{C}h\varepsilon(1 + |x|^2) + \frac{h}{\varepsilon}|\sigma(x)w|^2 + |\sigma(x)w|^2 \\
&\leq |x|^2 + 2h(\tilde{\beta} - \tilde{\alpha}|x|^2) + \tilde{C}(\varepsilon + h\varepsilon + h^2)(1 + |x|^2) + \left(1 + \frac{h+1}{\varepsilon}\right)|\sigma(x)w|^2.
\end{aligned}$$

Now, set $\varepsilon = h^2$ and choose h small enough such that $\tilde{C}(\varepsilon + h\varepsilon + h^2) \leq \tilde{\alpha}h$. Therefore,

$$|F_h(x, w)|^2 \leq |x|^2 + h(\tilde{\gamma} - \tilde{\alpha}|x|^2) + \left(1 + \frac{h+1}{\varepsilon}\right)|\sigma(x)w|^2$$

where $\tilde{\gamma} = 2\tilde{\beta} + \tilde{\alpha}$. Then

$$|F_h(x, w)|^2 \leq (1 - \tilde{\alpha}h)|x|^2 + h\gamma + \left(1 + \frac{h+1}{\varepsilon}\right)|\sigma(x)w|^2.$$

By assumption σ is a bounded function on \mathbb{R}^d . Then, there exists $\tilde{C} > 0$ depending on h and σ such that

$$|F_h(x, w)|^2 \leq (1 - \tilde{\alpha}h)|x|^2 + \tilde{C}(1 + |w|^2).$$

Using the classical inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, we finally get the existence of $\gamma \in (0, 1)$ and $C > 0$ such that for all $(x, w) \in \mathbb{R}^d \times \mathbb{R}^d$

$$|F_h(x, w)| \leq \gamma|x| + C(1 + |w|) \quad (2.12)$$

which achieves the proof of **(H₁)**.

We now turn to the proof of **(H₂)**. Let $K > 0$ and take $\mathbf{x} = (x, x', y, y') \in B(0, K)^4$. Here we take $\tilde{K} = K$. Hence, let us now define $\Lambda_{\mathbf{x}}$. For all $u \in B(0, K)$, we set

$$\Lambda_{\mathbf{x}}(u) = \sigma^{-1}(x')\sigma(x)u + \sigma^{-1}(x')(x - x' + h(b(x) - b(x'))) + \sigma^{-1}(x')\sigma(x)y - y' \quad (2.13)$$

which is equivalent to $\tilde{F}_h(x, u, y) = \tilde{F}_h(x', \Lambda_{\mathbf{x}}(u), y')$ for all $u \in B(0, K)$.

Hence, for all $u \in B(0, K)$,

$$J_{\Lambda_{\mathbf{x}}}(u) = \sigma^{-1}(x')\sigma(x). \quad (2.14)$$

Since σ, σ^{-1} and b are continuous, there exist $C_K > 0$ and $m_K > 0$ independent from \mathbf{x} such that for all $u \in B(0, K)$,

$$\begin{aligned}
|\det(J_{\Lambda_{\mathbf{x}}}(u))| &\geq C_K \\
|\Lambda_{\mathbf{x}}(u) - u| &\leq m_K.
\end{aligned}$$

Then, we extend $\Lambda_{\mathbf{x}}$ to a continuous and invertible function in such a way that there exists $M_K \geq m_K$ such that for all $u \in \mathbb{R}^d$,

$$|\Lambda_{\mathbf{x}}(u) - u| \leq M_K.$$

Finally, the function \tilde{F}_h satisfies **(H₂)**. □

The two following subsections are devoted to outline examples of sequences which satisfy hypothesis **(H_{exp})** or **(H_{poly})**. In particular, Subsection 2.6 includes the case where the process $(\Delta_n)_{n \in \mathbb{Z}}$ corresponds to fractional Brownian motion increments.

2.5 An explicit case which satisfies $(\mathbf{H}_{\text{exp}})$

In this subsection, we investigate an explicit exponential case with the following definition for the sequence $(a_k)_{k \geq 0}$

$$a_0 = 1 \quad \text{and} \quad \forall k \in \mathbb{N}^*, \quad a_k = C_a e^{-\lambda k} \quad (2.15)$$

with $C_a \in \mathbb{R}$. Let us recall that $b_0 = 1$ (since $a_0 = 1$) and for all $k \geq 1$, we can get the following general expression of b_k (see Appendix A):

$$b_k = \sum_{p=1}^k \frac{(-1)^p}{a_0^{p+1}} \left(\sum_{\substack{k_1, \dots, k_p \geq 1 \\ k_1 + \dots + k_p = k}} \prod_{i=1}^p a_{k_i} \right). \quad (2.16)$$

A classical combinatorial argument shows that $\sharp\{(k_1, \dots, k_p) \in \mathbb{N}^* \mid k_1 + \dots + k_p = k\} = \binom{k-1}{p-1}$. As a consequence, when the sequence $(a_k)_{k \geq 0}$ is defined by (2.15), we can easily compute the coefficients b_k for $k \geq 1$,

$$\begin{aligned} b_k &= \sum_{p=1}^k (-C_a)^p e^{-\lambda k} \sharp\{(k_1, \dots, k_p) \in \mathbb{N}^* \mid k_1 + \dots + k_p = k\} \\ &= \sum_{p=1}^k \binom{k-1}{p-1} (-C_a)^p e^{-\lambda k} \\ b_k &= -C_a (1 - C_a)^{k-1} e^{-\lambda k}. \end{aligned} \quad (2.17)$$

Hence, to satisfy $(\mathbf{H}_{\text{exp}})$, we only need C_a to be such that $\mu := \lambda - \ln |1 - C_a| > 0$ and then for all $k \in \mathbb{N}^*$, we get

$$|b_k| \leq C_b e^{-\mu k} \quad (2.18)$$

with $C_b > 0$ a constant depending on C_a .

Remark 2.8. \triangleright In this setting where everything is computable, it's interesting to see that the asymptotic decrease of the sequence $(|b_k|)$ is not only related to the one of the sequence $(|a_k|)$. For instance, if we take $C_a < 0$, the simple fact that $a_0 > 0$ and $a_k < 0$ for all $k > 0$ makes (b_k) diverge to $+\infty$ and nevertheless, $(|a_k|)$ decreases to 0 at an exponential rate.

\triangleright If we take $C_a = 1$, we can reduce $(\Delta_n)_{n \in \mathbb{Z}}$ to the following induction

$$\forall n \in \mathbb{Z}, \quad \Delta_{n+1} = \xi_{n+1-k} + e^{-\lambda} \Delta_n. \quad (2.19)$$

2.6 Stationary Gaussian sequence of fractional type

Let $H \in (0, 1)$. In the sequel, we will speak about stationary Gaussian sequence of *fractional type* if the sequence (a_k) is such that

$$\forall k \geq 0, \quad |a_k| \leq C_\rho (k+1)^{-\rho} \quad \text{and} \quad |a_k - a_{k+1}| \leq \tilde{C}_\rho (k+1)^{-(\rho+1)} \quad (2.20)$$

with $\rho := 3/2 - H$.

As we will see below, this condition includes the case where $(\Delta_n)_{n \in \mathbb{Z}}$ corresponds to the fractional Brownian motion (fBm) increments. Unfortunately, computing the rate of convergence of the corresponding sequence $(b_k)_{k \geq 0}$ is a hard task and strongly depends on the variations of $(a_k)_{k \geq 0}$. Actually, in Proposition 2.2 and 2.3, dealing with the same order of memory, we will see that the orders of rate of convergence are really different. Note that the first corresponds to the case where $a_k := (k+1)^{-(3/2-H)}$ for all $k \in \mathbb{N}$ whereas the second deals with fBm increments.

Proposition 2.2. Assume (\mathbf{H}_1) and (\mathbf{H}_2) . Let $H \in (0, 1)$ and set $\rho := 3/2 - H \in (1, 3/2)$. If for all $k \geq 0$, $a_k = (k+1)^{-\rho}$, then (a_k) is of fractional type and we have $|b_k| \leq (k+1)^{-\rho}$. Moreover, if $\rho > 3/4$ Theorem 2.1 (ii) holds with the rate

$$v(\rho, \rho) = v(3/2 - H, 3/2 - H) = \frac{1}{2} \begin{cases} (1-H)^2 & \text{if } H \in (0, 1/2] \\ (3-4H)^2 & \text{if } H \in (1/2, 3/4). \end{cases}$$

Remark 2.9. This result follows from the proof of the inequality $|b_k| \leq (k+1)^{-\rho}$ for all $k \geq 0$ which is outlined in Appendix B. The key argument in this proof is the property of log-convexity of the sequence $(a_k)_{k \in \mathbb{N}}$, which means that for all $k \in \mathbb{N}$, $a_k \geq 0$ and for $k \geq 1$, $a_k^2 - a_{k-1}a_{k+1} \leq 0$.

As mentioned before, the terminology “fractional type” refers to the fractional Brownian motion. Indeed, in a continuous-time setting, a famous and classical example of non-Markovian dynamics is SDE driven by fBm

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t \quad (2.21)$$

Recall that a d -dimensional fBm with Hurst parameter $H \in (0, 1)$ is a centered Gaussian process $(B_t)_{t \geq 0}$ with stationary increments satisfying

$$\forall t, s \geq 0, \forall i, j \in \{1, \dots, d\}, \quad \mathbb{E} \left[(B_t^i - B_s^i)(B_t^j - B_s^j) \right] = \delta_{ij} |t - s|^{2H}.$$

The study by a coupling argument of the rate of convergence to equilibrium for this kind of dynamics has been undertaken by Hairer [14], Fontbona and Panloup [11], Deya, Panloup and Tindel [9], respectively in the additive noise, multiplicative noise with $H > 1/2$ and multiplicative noise with $H \in (1/3, 1/2)$. Here in our discrete-time setting, we are thus concerned by the long time behavior of (2.1) if we take for $h > 0$

$$(\Delta_n)_{n \in \mathbb{Z}} = (B_{nh} - B_{(n-1)h})_{n \in \mathbb{Z}} \quad (2.22)$$

which is a stationary Gaussian sequence. It can be realized through a moving average representation with coefficients $(a_k^H)_{k \geq 0}$ defined by (see [19]):

$$a_0^H = h^H \kappa(H) 2^{1/2-H} \quad \text{and for } k \geq 1, \quad a_k^H = h^H \kappa(H) \left(\left(k + \frac{1}{2} \right)^{H-1/2} - \left(k - \frac{1}{2} \right)^{H-1/2} \right) \quad (2.23)$$

where

$$\kappa(H) = \frac{\sqrt{\sin(\pi H) \Gamma(2H+1)}}{\Gamma(H+1/2)}.$$

One can easily check that $a_k^H \underset{k \rightarrow +\infty}{\sim} C_{h,H} (k+1)^{-(3/2-H)}$ and $|a_k^H - a_{k+1}^H| \leq C'_{h,H} (k+1)^{-(5/2-H)}$.

Hence $(a_k^H)_{k \geq 0}$ is of fractional type in the sense of (2.20). Now, the question is: how does the corresponding (b_k^H) behave? When H belongs to $(0, 1/2)$, only a_0^H is non-negative and then (a_k^H) is not log-convex. Therefore, we cannot use this property to get the asymptotic behavior of (b_k^H) as we did in Proposition 2.2. However, thanks to simulations (see Figure 1a and 1b), we conjectured and we proved (see Appendix E) the following proposition.

Proposition 2.3. There exists $C''_{h,H} > 0$ such that for all $H \in (0, 1/2)$

$$\forall k \geq 0, \quad |b_k^H| \leq C''_{h,H} (k+1)^{-(H+1/2)}. \quad (2.24)$$

Then, if we assume (\mathbf{H}_1) and (\mathbf{H}_2) , Theorem 2.1 (ii) holds with the rate

$$v(\rho, 2-\rho) = v(3/2 - H, H + 1/2) = \frac{1}{2} \begin{cases} H(1-2H) & \text{if } H \in (0, 1/4] \\ \frac{1}{8} & \text{if } H \in (1/4, 1/2). \end{cases}$$