



HAL
open science

A general martingale approach to large noise homogenization

Dimitri Faure, Mathias Rousset

► **To cite this version:**

Dimitri Faure, Mathias Rousset. A general martingale approach to large noise homogenization. *Electronic Journal of Probability*, 2024, 29, pp.1-49. 10.1214/24-EJP1177 . hal-04363393v2

HAL Id: hal-04363393

<https://hal.science/hal-04363393v2>

Submitted on 20 Dec 2024

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.



Distributed under a Creative Commons Attribution 4.0 International License

A general martingale approach to large noise homogenization

Dimitri Faure* Mathias Rousset†

Abstract

We consider Markov processes with generators of the general form $\gamma\mathcal{L}_1 + \mathcal{L}_0$, in which \mathcal{L}_1 is associated to a so-called *dominant process*. The dominant process is assumed to converge for large times towards a random point in a given subset of the state space called the *effective state space*. Using the usual characterization through martingale problems, we give general conditions under which the following homogenization theorem holds true: the original process converges in law, when γ is large and for both the (Meyer-Zheng) pseudo-path topology and finite-dimensional time marginals, towards an identified effective Markov process defined on the effective state space. Few simple model examples for diffusions are studied.

Keywords: Markov processes; Homogenization; Large noise; Quantum continuous measurements.

MSC2020 subject classifications: 60J; 60F05.

Submitted to EJP on -, final version accepted on -.

Supersedes arXiv:2304.09624.

Contents

1	Introduction	2
2	Setting, assumptions and main theorems	7
2.1	Notation	7
2.2	Setting and assumptions	8
2.3	The case of diffusions with Lipschitz coefficients	12
2.4	Main homogenization theorem	13

*DMA, École normale supérieure, Université PSL, CNRS, 75005 Paris, France.

†IRMAR and Inria, University of Rennes, France.

3 Proof of the main theorems	14
3.1 Step 1: Tightness and extraction	14
3.2 Step 2: The limit process lies in the effective space	15
3.3 Step 3: The limit is càdlàg	17
3.4 Step 4: The limit is the unique solution of a martingale problem	20
3.5 Proof of Corollary 2.11	21
3.6 Proof of Proposition 2.4	23
4 Examples	29
4.1 Diffusion in a simplex	29
4.2 Diffusion in a strip	31
4.3 Diffusion in an Euclidean ball	34
Appendices	36
A The pseudo-paths topology and the Meyer-Zheng criterion	36
A.1 The pseudo-path topology	36
A.2 Convergence of finite-dimensional distributions	37
A.3 Compact sub-spaces	38
A.4 The Meyer-Zheng criterion	38
A.5 Characterization of càdlàg martingales	39
B Some remarks on diffusions	39
C Calculus associated with the diffusion in a strip	40
D Calculus associated with the diffusion in the Euclidean ball	46
References	48

1 Introduction

A formalism for strong noise homogenization. Let E be a Polish state space, and let $E_{\text{eff}} \subset E$ be a closed subset, hereafter called the *effective state space*. This paper considers on E a family of (càdlàg) Markov processes denoted $(X_t^\gamma)_{t \geq 0} := (X_t^{(\gamma,1)})_{t \geq 0}$ with (formal) infinitesimal Markov generator of the form¹

$$\gamma \mathcal{L}_1 + \mathcal{L}_0, \tag{1.1}$$

in which $\mathcal{L}_1, \mathcal{L}_0$ are two (formal) infinitesimal Markov generators, and $\gamma > 0$ is a positive parameter that is meant to go to infinity ($\gamma \rightarrow +\infty$). This process will be called here the *initial process*.

In this setting, the following main structural assumption lay the basis to what is here called *large noise homogenization*. For each initial condition $x \in E$, the process generated by \mathcal{L}_1 (called here after the *dominant process*) and denoted $(X_t^{(1,0)}(x))_{t \geq 0}$, is assumed to converge in probability for large time towards a random point in E_{eff} :

$$\lim_{t \rightarrow +\infty} X_t^{(1,0)}(x) = X_\infty^{(1,0)}(x) \in E_{\text{eff}} \quad \text{in probability,} \tag{1.2}$$

whose distribution will be denoted:

$$\mathcal{P}(x, \cdot) := \text{Law}(X_\infty^{(1,0)}(x)).$$

¹One can equivalently consider the small perturbation problem $\mathcal{L}_1 + \frac{1}{\gamma} \mathcal{L}_0$ on large time scales of order γ .

One also wants to define E_{eff} as the smallest effective space satisfying (1.2), which suggests the following associated additional assumption:

$$\mathcal{P}(x, \cdot) = \delta_x, \quad \forall x \in E_{\text{eff}}. \tag{1.3}$$

The property (1.2) also implies by definition of \mathcal{P} and by the Markov property that $\mathcal{L}_1 \mathcal{P}\varphi = 0$ on E for any continuous bounded test function $\varphi : E_{\text{eff}} \rightarrow \mathbb{R}$ defined on the effective state space. One can then define the operator

$$\mathcal{L}_{\text{eff}}\varphi := \mathcal{L}_0 \mathcal{P}\varphi = (\gamma \mathcal{L}_1 + \mathcal{L}_0) \mathcal{P}\varphi, \tag{1.4}$$

which is (formally) a Markov generator on E_{eff} as the formal limit when $s \rightarrow 0$ of pure jump Markov generators on E_{eff} :

$$\mathcal{L}_{\text{eff}} = \lim_{s \rightarrow 0^+} (e^{s\mathcal{L}_0} \mathcal{P} - \text{Id}) / s = \lim_{s \rightarrow 0^+} (e^{s\mathcal{L}_0} \mathcal{P} - \mathcal{P}) / s;$$

in which we have used the ‘minimality’ property (1.3) of E_{eff} which states that $\mathcal{P} = \text{Id}$ as operators on functions of E_{eff} . By (1.4), the process

$$t \mapsto \mathcal{P}\varphi(X_t^\gamma) - \int_0^t \mathcal{L}_0 \mathcal{P}\varphi(X_s^\gamma) ds \tag{1.5}$$

is expected to be a martingale for the natural filtration generated by X^γ and for all φ in a class of sufficiently regular bounded test functions. Since by (1.2) the dominant generator \mathcal{L}_1 tends to confine the process X^γ near E_{eff} when γ is large, one may then expect the following two facts:

- i) X^γ behaves when γ is large approximately as a Markov process taking values in E_{eff} and called hereafter the *effective process*,
- ii) this effective process is a solution to the martingale problem in E_{eff} associated with \mathcal{L}_{eff} (by (1.5) and the fact that $\mathcal{P}\varphi = \varphi$ on E_{eff}).

Making the previous two statements completely rigorous and identifying the effective process as the unique solution of a martingale problem is precisely what we call here the *martingale problem approach to strong noise homogenization*. This will be carried out in the present work in a quite general setting using (as already proposed in previous works mentioned below) the pseudo-path topology studied in the seminal work [18] of Meyer and Zheng. This should be contrasted with e.g. [20], a classical reference on the martingale problem approach in a classical homogenization setting, which is using the Skorokhod (‘uniform’) topology.

Note that, in strong noise problems, even if the initial process is a diffusion, the effective process obtained as a limit may be of jump type. As a consequence, convergence with respect to the stronger Skorokhod topology (instead of the pseudo-path topology) cannot be true.

A class of examples for diffusions (see Section 2.3 and Section 4). The main examples studied in this work (Section 4) will be of the following form: E will be a closed subset of a manifold (say \mathbb{R}^d), E_{eff} will be a smooth sub-manifold of E included in the boundary ∂E of E , and $t \mapsto X_t^{\gamma,1}$ will be a diffusion. The dominant process will be a diffusion which eventually almost surely hits E_{eff} (albeit in our examples for an infinite hitting time); and the perturbation process will be a deterministic flow with a non-vanishing drift at E_{eff} (pointing inward at the boundary ∂E when $E_{\text{eff}} \subset \partial E$).

Note that in the context of potential theory, if the dominant process is a (singularly at E_{eff}) time-changed Brownian motion and $E_{\text{eff}} = \partial E$ is the boundary of a smooth

domain E (see Section 4.3), then the kernel $\mathcal{P}(x, \cdot)$ is exactly the so-called *harmonic measure* of E with pole at x . When \mathcal{L}_0 is an inward vector field, the Markovian effective generator (1.4) is also known as a quite standard general version of the so-called *Dirichlet-to-Neumann* operator. The latter has been extensively studied in the inverse problem literature (see *e.g.* the seminal work [25]) for reasons of independent interest. The Dirichlet-to-Neumann effective operator is, in fact, a non-diffusive Markov operator on E_{eff} of Levy type, as will be demonstrated in the examples of Section 4, especially the example of Section 4.3. More comments on that topic will be given in the last paragraph of this introduction.

Previous work. Motivated by quantum continuous measurements (see [7, 14, 1, 19, 26]), or filtering theory (see [4]), strong noise homogenization has recently been rigorously studied in the special case where E_{eff} is a finite set.

In [2], the authors show strong noise homogenization for a certain class of diffusion processes having a generator of the form² (1.1). In order to prove their theorem, the authors develop a *finite-dimensional* homogenization theorem [2, Theorem 3.1] for *bounded operators* and, due to the specific form of the diffusion processes considered (linear drift and quadratic mobility), are able to prove convergence by developing a perturbative argument.

In [3], the homogenization result is refined by incorporating a decoration with an interval-valued process describing the limit of the "spikes" of the initial process.

In [13], the author adapts the proof of [2, Theorem 3.1] for diffusion processes living in a simplex $K \subset \mathbb{R}^n$ whose vertices define E_{eff} , and with a dominant process of martingale type. This ensures that the kernel of \mathcal{L}_1 is the finite-dimensional subspace of affine functions of \mathbb{R}^d which simplifies the analysis.

More comments comparing our result with [13] and [3] will be given below, in the next two paragraphs of this introduction.

All the above works consider the so-called (Meyer-Zheng) pseudo-path topology (see [18] and Section A) on trajectories, as well as the associated Meyer-Zheng "tightness" criterion, in order to handle strong noise homogenization. The use of the Meyer-Zheng criterion in a homogenization context is however neither new nor limited to the recent consideration of strong noise problems. It has been quite consistently used more than a decade ago to handle other intricate cases, for instance periodic problems (see [21]), or backward stochastic differential equations (see [22]).

New results. By considering the classical martingale problem approach, this paper generalizes the previously mentioned body of work. The basic idea consists in remarking that the pseudo-path topology and the associated Meyer-Zheng criterion are well-suited to martingale problems, which is a standard tool for weak convergence (*i.e.* convergence in law) results in classical homogenization of stochastic dynamics. Note that in classical references (such as [20]), limits of processes are handled with the Skorokhod topology, which is not possible here (see [11] for a general reference on convergence of processes and martingale problems studied under the Skorokhod topology).

In the present work, the main homogenization theorem is obtained by making the following four steps rigorous.

First, i) relative compactness of trajectory distributions is, by definition of the pseudo-path topology, a direct consequence of Prokhorov theorem with an appropriate containment condition in a compact set. One can then consider extracted converging sequences.

²More precisely, in [2] a very slightly more general situation is considered with the possible additional presence of a generator at an intermediate time scale (in terms of γ).

Second, ii) one needs to check that the limiting pseudo-path takes its values in E_{eff} . This step may be quite technical, and we propose a relatively natural setting that enables to carry it out by first proving an intermediate property: the dominant process converges towards E_{eff} for large times uniformly with respect to the initial condition taken in a compact set (see Proposition 2.4).

Third, iii) càdlàg processes (in fact quasi-martingales) constructed from the full process (1.1) with appropriate test functions of the form $\mathcal{P}\varphi$ are expected to have by (1.4) a mean variation uniformly bounded in γ : this is the so-called Meyer-Zheng criterion. This criterion (a key result of [18]) ensures that the large γ limit in the pseudo-path topology of such processes inherits the càdlàg regularity and the quasi-martingale property. We stress that without the Meyer-Zheng criterion, such limits may be too degenerate to handle the martingale property. With this guarantee, it is then possible to show that the extracted sequence of processes in Step i) does satisfy the martingale problem associated with \mathcal{L}_{eff} .

Finally, iv) one simply needs to assume that the obtained effective martingale problem is well-posed (there is a unique càdlàg process distribution solution of it). This ends the proof of convergence in distribution of the considered process towards the effective process in the pseudo-path topology. Extension of the convergence to finite-dimensional time-marginals are also obtained.

In order to make the above four steps i), ii), iii) and iv) rigorous, the following two assumptions will be especially important. The first one (later on denoted Assumption 2) is the continuity of the dominant process hitting distribution $\mathcal{P}(x, \cdot)$ with respect to the initial condition x . This assumption is key in proving the large time convergence of the dominant process uniformly with respect to the initial condition, which is instrumental in Step ii). It is inherited from [13] in which it is argued that it is somehow unavoidable using a *pathological* counter-example ([13, Section 4]). The second important assumption is of course the fact that the effective generator \mathcal{L}_{eff} is well-defined, continuous near E_{eff} , and defines a well-posed martingale problem. Well-posedness of martingale problems may not be so easy to obtain, but it is a topic on its own (well covered at least in \mathbb{R}^d) in the literature. When the effective process can be described by a stochastic differential equation which is itself well-posed in the strong sense (e.g. with Lipschitz coefficients), quite generic results exist (see e.g. the reference [17]).

Let us finally detail the difference between our approach and the one adopted in [13]. In the latter, the homogenization problem is specific (see details in Section 4.1): the dominant process is a martingale and the space of harmonic functions (the kernel of \mathcal{L}_1) can be identified with the space of affine functions on \mathbb{R}^n . The main idea of [13] is then to couple the initial process with an approximate process whose generator preserves affine functions and can be homogenized *directly for any given finite* γ . However, this strategy only works for this specific case. In contrast, the approach based on characterization by martingale problems developed in the present paper is independent of the specific properties of the harmonic functions, and allows us to prove general homogenization results.

A simple example of a new limit theorem associated with the phenomenology of "spikes". As an example of new result obtained in this paper, one may consult the example of Section 4.2. There, one considers an already studied one dimensional toy model for strong noise homogenization which consists of a dominant time-changed Brownian motion in $[-1, 1]$ perturbed by a drift. Previous works (see [2, 13]) have shown the convergence towards a jump process in $E_{\text{eff}} = \{+1, -1\}$. The novelty consists in enriching the process with its quadratic variation, leading to a process defined on $E = [-1, 1] \times \mathbb{R}_+$. This enrichment may be thought of as a way to keep track of the

"size" of the trajectories as measured by quadratic variation. We prove convergence of the enriched process to a limiting effective process in $\{-1, +1\} \times \mathbb{R}_+$. The latter is the sum of two generators.

The first is a jump generator creating simultaneously a switching between $+1$ and -1 for the first variable and a positive jump in \mathbb{R}_+ for the second variable (the associated intensity and distribution are proportional to a given explicit finite measure denoted ν in Section 4.2).

The second is a Levy generator in \mathbb{R}_+ of subordinator type for the second variable (the first variable in $\{-1, +1\}$ stays constant). The associated Levy measure is proportional to:

$$\mu(dh) := \left(\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 e^{-h \frac{\pi^2 k^2}{8}} \right) dh.$$

Note that the jumps of this subordinator is a mixture of exponentially distributed jumps of size $\frac{8}{\pi^2 k^2}$ with $k \in \mathbb{Z}$.

The jumps of the effective process of this example can be interpreted as the "sizes" (measured with quadratic variation) of the different "spikes" of the original process inside the interval $] - 1, +1[$. This phenomenology of "spikes" and its physical relevance has already been pointed out and discussed in [26] in the context of continuous measurements of a quantum system. Mathematical analysis has then followed for instance in [2] and [3]. Our approach of the above toy example is somehow complementary to the one of [3]. In the latter reference, the authors directly analyse the limit of the "spikes" as an interval-valued process decorating the effective jump process, while we incorporate the "size" of the "spikes" (measured with quadratic variation) as an additional variable of the effective, homogenized process.

Link with the boundary process associated with reflected diffusions. Finally, we want to suggest some potentially interesting connections between strong noise homogenization and reflected diffusions defined as solutions of the Skorokhod reflection problem (see e.g. [23] for a general reference on the topic). The context is the following: consider a regular diffusion with generator $\tilde{\mathcal{L}}_1$ which hits E_{eff} at a finite (as opposed to typically infinite) time. The latter and the dominant process considered in the present work can usually be obtained from one another using a time-change scalar field $a > 0$ diverging near E_{eff} : $\tilde{\mathcal{L}}_1 = a \times \mathcal{L}_1$. In addition $E_{\text{eff}} = \partial E$ is assumed to be the smooth boundary of regular domain E of, say, \mathbb{R}^n , and the perturbation generator $\mathcal{L}_0 = F_{(0)}^i \partial_i$ is a vector field which is inward-pointing at the boundary ∂E . We have already mentioned above that in this context the effective process generator (1.4) is the so-called *Dirichlet-To-Neumann* (Markov) operator defined on smooth functions of E_{eff} , which has been studied in the inverse problem literature for linear elliptic partial differential equations (see e.g. the seminal work [25]). It is quite interesting to remark that in [15], the author gives a probabilistic (and completely different) interpretation of \mathcal{L}_{eff} using reflected diffusions.

Let us briefly describe the idea. Consider the diffusion constrained in E by the reflection at the boundary ∂E of E generated by the inward vector field $F_{(0)}$. The latter (see [23]), denoted here $(X_t)_{t \geq 0}$, is classically defined as the solution of the so-called Skorokhod problem, which is obtained by adding a term to the stochastic differential equation satisfied by the non-reflected diffusion. This additional term is of the form $F_{(0)}(X_t)dL_t$ where L_t is a non-decreasing adapted continuous process that is constant outside the boundary and must satisfies $dL_t = \mathbf{1}_{X_t \in \partial E} dL_t$. The process L_t is a kind of 'local time' that has to be determined as an output of the Skorokhod problem. The *boundary process* is then the Markov process $l \mapsto X_l^{\partial E} = X_{\tau_l}$ obtained by indexing

the initial process with this ‘local time’ using the right continuous inverse defined by the stopping time $\tau_l := \inf(t : L_t > l)$. If $\mathcal{P}(x, \cdot)$ denotes the hitting distribution at the boundary of E of the underlying non-reflected diffusion \mathcal{L}_1 , it can be shown using a simple application of Itô formula to $\mathcal{P}(\varphi)$ for φ smooth (see [15] for the special case of Brownian motion and normal reflection) that the boundary process $(X_{\tau_l})_{l \geq 0}$ is precisely a solution of the martingale problem associated with $\mathcal{L}_{\text{eff}} = \mathcal{L}_0 \mathcal{P}$. The latter can thus be interpreted as the associated infinitesimal generator.

Developing the connections between strong noise homogenization and boundary processes lies outside the scope of this paper and is left as an open problem of interest. We simply want to mention that reflected diffusions may also be obtained as a limit of diffusions of the form (1.1) using time-changes of the dominant and/or the perturbation process. It is thus not completely surprising to obtain the same ‘trace’ at the boundary.

Note also that in dimension one, this connection has been used in depth in the proof of the spike-enhanced strong noise homogenization result obtained in [3]. In the latter reference, the effective two-points jump process is decorated with the limit of the graph of the spikes of the initial process. In spirit, the proof relies on a coupling argument with a (reflected) time-changed Brownian motion. The change in time indeed becomes in the limit the inverse of a local time at the boundary of the considered interval. The spikes are analyzed using Itô’s excursion theory.

Organization. This work is divided as follows. In Section 2, we rigorously introduce notation, then state the different assumptions, and finally state the main strong noise homogenization theorems. In Section 3, the main theorems are proved. Finally in Section 4, explicit examples are described and are studied in detail.

2 Setting, assumptions and main theorems

2.1 Notation

- E main Polish state space with metric d .
- $(X_t^{(\gamma_1, \gamma_0)})_{t \geq 0}$: a Markov process with generator $\gamma_1 \mathcal{L}_1 + \gamma_0 \mathcal{L}_0$ defined on E .
- $X^\gamma := X^{(\gamma, 1)}$: a shorthand notation.
- $\mu_0^\gamma := \text{Law}(X_0^\gamma)$: a given family of initial distributions in E indexed by γ .
- $t \mapsto X_t(x)$: a Markov process with explicitly mentioned initial condition $X_0(x) = x$.
- \mathbb{E}_x : the expectation when considering a Markov process $(X_t(x))_{t \geq 0}$ with initial condition $x \in E$.
- \mathbb{E}_μ : the expectation when considering a Markov process $(X_t)_{t \geq 0}$ with initial distribution $\text{Law}(X_0) = \mu$.
- $L^0(e^{-t} dt, E)$: the equivalence class of measurable trajectories $\mathbb{R}_+ \rightarrow E$ that are Lebesgue almost everywhere equal. It is endowed with the (Polish) pseudo-path topology.
- $\text{Pr}(E)$: the space of probability distributions on E .
- $B(x, r)$: a closed ball of center x and radius r
- $x \cdot y = \langle x, y \rangle$: the two notations may be used for scalar product in Euclidean space.

2.2 Setting and assumptions

Let E denote a Polish state space. For each non-negative $\gamma_1, \gamma_0 \geq 0$, we assume that we can measurably associate to any initial condition $x \in E$ a càdlàg Markov process denoted

$$t \mapsto X_t^{(\gamma_1, \gamma_0)}(x),$$

whose (formal) infinitesimal generator is of the form

$$\gamma_1 \mathcal{L}_1 + \gamma_0 \mathcal{L}_0.$$

In the above, \mathcal{L}_1 (resp. \mathcal{L}_0) is the generator associated with the so-called *dominant* $X^{(1,0)}$ (resp. *perturbation* $X^{(0,1)}$) process.

In what follows, we denote by

$$\mu_0^\gamma = \text{Law}(X_0^\gamma)$$

the initial probability distribution of the considered process. Throughout this paper we will assume that μ_0^γ converges when γ is large to a limit μ_0 in $\text{Pr}(E)$.

The main structural assumption underlying the present work is the following: for any initial condition $x \in E$, the dominant process $X_t^{(1,0)}(x)$ converges for large times t towards a random variable $X_\infty^{(1,0)}(x)$ taking values in a closed subset $E_{\text{eff}} \subset E$ (at least in probability, although the almost sure convergence is typical). More precisely:

Assumption 1 (Main assumption). *For each $x \in E$, there is a random variable denoted $X_\infty^{(1,0)}(x) \in E_{\text{eff}}$ taking values in a closed subset $E_{\text{eff}} \subset E$ such that, in probability,*

$$\lim_{t \rightarrow +\infty} X_t^{(1,0)}(x) = X_\infty^{(1,0)}(x) \in E_{\text{eff}}.$$

Moreover, E_{eff} is minimal in the sense that if $x \in E_{\text{eff}}$, then $X_\infty^{(1,0)}(x) = x$.

The limiting distribution thus defines a measurable probability kernel denoted

$$\mathcal{P}(x, \cdot) := \text{Law}(X_\infty^{(1,0)}(x)) \in \text{Pr}(E_{\text{eff}}).$$

As explained in the introduction, the main goal of this work is to study the limit distribution of the process $X^{(\gamma,1)}$ when $\gamma \rightarrow +\infty$, and to show that the latter is a Markov process on E_{eff} with infinitesimal generator given by

$$\mathcal{L}_{\text{eff}}[\varphi] := \mathcal{L}_0[\mathcal{P}[\varphi]].$$

Remark 2.1 (On invariances by time-changes). First remark: by definition, the process $t \mapsto X_{t/k}^{(k\gamma, k\gamma_0)}$ has a generator independent of k so that its distribution can be assumed to be independent of $k > 0$. Second remark: the effective generator \mathcal{L}_{eff} depends on the dominant generator only through the probability kernel \mathcal{P} . As a consequence, the modified strong noise problem associated with $\gamma\rho\mathcal{L}_1 + \mathcal{L}_0$ where $\rho : E \rightarrow]0, +\infty[$ is a function bounded away from 0 will lead to the same effective process (indeed the action of the field $\rho > 0$ of the dominant generator can be interpreted as a time-change of the dominant process, which results by definition in a kernel \mathcal{P} independent of ρ).

Our first technical assumption amounts to assume that \mathcal{P} is continuous:

Assumption 2 (Continuity of \mathcal{P}). *The map $x \in E \mapsto \mathcal{P}(x, \cdot) \in \text{Pr}(E_{\text{eff}})$ is continuous, $\text{Pr}(E_{\text{eff}})$ being endowed with the usual Polish topology of convergence in distribution.*

In particular, since by assumption, for all $x_0 \in E_{\text{eff}}$, $\mathcal{P}(x_0, \cdot) = \delta_{x_0}$, we obviously have $\lim_{x \rightarrow x_0} \mathcal{P}(x, \cdot) = \delta_{x_0}$ for all $x_0 \in E_{\text{eff}}$.

The next assumption defines a natural notion of containment in a compact set. The latter is sufficient in order to obtain the tightness of the considered sequence of processes. This condition is not a necessary condition for tightness; it might be relaxed in specific cases. Although it only has to be checked for $\gamma_0 \in \{0, 1\}$, it is typically true for any $\gamma_0 \in \mathbb{R}$.

Assumption 3 (Containment in a compact set). *Let $\gamma_0 \in \{0, 1\}$ be given. For any horizon time T , any $\varepsilon > 0$ and any compact $K \subset E$, there exists a compact set $K_{\varepsilon, T} \subset E$ such that:*

$$\sup_{x \in K, \gamma > 0} \mathbb{P}_x \left[\exists t \in [0, T], X_t^{(\gamma, \gamma_0)} \in K_{\varepsilon, T}^c \right] \leq \varepsilon. \tag{2.1}$$

Using a time scaling argument, it is possible to show the following direct consequence on the dominant process.

Lemma 2.2. *Assumption 3 implies that for any $\varepsilon > 0$ and any compact $K \subset E$, there exists a compact set $K_\varepsilon \subset E$ such that:*

$$\sup_{x \in K} \mathbb{P}_x \left[\exists t \in \mathbb{R}_+, X_t^{(1, 0)} \in K_\varepsilon^c \right] \leq \varepsilon. \tag{2.2}$$

Proof. Let $T = 1$ be fixed. By a time rescaling argument and routine monotone convergence, it holds

$$\begin{aligned} & \sup_{\gamma > 0} \mathbb{P}_x \left[\exists t \in [0, 1], X_t^{(\gamma, 0)} \in K_\varepsilon^c \right] \\ &= \sup_{\gamma > 0} \mathbb{P}_x \left[\exists t \in [0, \gamma], X_t^{(1, 0)} \in K_\varepsilon^c \right] = \mathbb{P}_x \left[\exists t \in \mathbb{R}_+, X_t^{(1, 0)} \in K_\varepsilon^c \right]. \end{aligned}$$

Assumption 3 then yields the result. □

Note also that since we consider a converging sequence of initial probability distributions $\mu_0^\gamma \rightarrow_{\gamma \rightarrow +\infty} \mu_0$ (a tight family of distributions), Assumption 3 implies by definition of tightness the existence of a compact $K_{\varepsilon, T}$ (using slightly abusively the same notation as in (2.1)) such that

$$\sup_{\gamma > 0} \mathbb{P}_{\mu_0^\gamma} \left[\exists t \in [0, T], X_t^{(\gamma, \gamma_0)} \in K_{\varepsilon, T}^c \right] \leq \varepsilon. \tag{2.3}$$

The next two assumptions (Assumptions 4 and 5) will be used as sufficient conditions in order to obtain a (technical) result given in Proposition 2.4 below, which states that the large time convergence of the dominant process towards E_{eff} is uniform with respect to its initial condition when taken in a given compact set.

The first assumption requires that the dominant process can be represented as a stochastic flow continuous with respect to the considered initial condition. By a stochastic flow representation, we mean that there exists a measurable map

$$(x, t, \omega) \mapsto X_t^{(1, 0)}(x)[\omega]$$

defined on $E \times \mathbb{R}_+ \times \Omega$ where $(\Omega, \mathcal{F}, \mathbb{P})$ denotes the underlying probability space and such that $t \mapsto X_t^{(1, 0)}(x)$ is distributed according to the dominant process with initial condition $X_0^{(1, 0)}(x) = x$. This happens typically if the dominant process is a strong solution of a Stochastic Differential Equation parametrized by the initial condition.

Assumption 4 (Continuous stochastic flow). *The dominant process can be represented by a stochastic flow such that, for each $t \geq 0$, $x \mapsto X_t^{(1, 0)}(x)$ is continuous in probability.*

The second assumption is a technical condition necessary in order to obtain the uniform convergence of Proposition 2.4 (see Remark 2.5 below for a counter-example).

Assumption 5 (Randomness of the hitting point). *The hitting point of the dominant process, $X_\infty^{(1,0)}(x)$, is deterministic if and only if $x \in E_{\text{eff}}$. In the latter case ($x \in E_{\text{eff}}$), we recall that $X_\infty^{(1,0)}(x) = x$.*

Remark 2.3 (On the role of the last two assumptions). Contrary to other assumptions, Assumption 4 and 5 are only used for the proof of Proposition 2.4 below. In the same way, Proposition 2.4 is only required once in Step 2 (that states that the limit effective process lies in E_{eff}) of the proof of the main homogenization theorem.

Proposition 2.4 (The large time convergence of the dominant process is uniform). *Assume that the dominant process satisfies Assumption 1-2-3-4-5 (Nota Bene: for the dominant process Assumption 3 amounts to (2.2)). Then there exists a continuous bounded non-negative function f with $f > 0$ on $E \setminus E_{\text{eff}}$ and $f = 0$ on E_{eff} such that for any compact $K \subset E$:*

$$\lim_{t \rightarrow +\infty} \sup_{x \in K} \mathbb{E} \left[f(X_t^{(1,0)}(x)) \right] = 0.$$

The proof of this result is quite technical and is postponed to the end of Section 3.

Remark 2.5 (Assumption 5 is necessary). The only unusual requirement of Proposition 2.4 is Assumption 5. It is however unavoidable. Indeed, consider the deterministic process on the unit circle given by the vector field $\sin^2(\frac{\theta}{2}) \cdot \partial_\theta$, where $\theta \in]0, 2\pi[$. The process evolves counter-clockwise (increasing θ), and eventually converges to $\theta = 2\pi = 0$ in infinite time. However, the hitting time of some $\theta_1 \neq 0$ starting with initial condition θ_0 becomes infinite when $\theta_0 \rightarrow 0^+$: if we start at θ_0 just above 0, we reach $\theta_1 > 0$ only after an arbitrarily large time. As a consequence, the process cannot be arbitrarily close to 0 at a large given time uniformly in the starting point; and the conclusion of Proposition 2.4 cannot hold true.

The next assumption is a usually benign continuity assumption which ensures that the addition of a small perturbation process results in a small change of the process distribution.

Assumption 6 (Continuous perturbation). *Let φ be any continuous bounded function. One has the continuity property: for all $t \geq 0$ and all $K \subset E$ compact,*

$$\lim_{\gamma \rightarrow +\infty} \sup_{x \in K} \left| \mathbb{E}(\varphi(X_t^{(1,1/\gamma)}(x))) - \mathbb{E}(\varphi(X_t^{(1,0)}(x))) \right| = 0$$

We can finally state important assumptions on the limit process, its generator, and the associated martingale problem (Assumption 7 and 8 below). Assumption 7 is a technical requirement that ensures that an effective generator indeed exists, with appropriate continuity properties.

Assumption 7 (Existence of an effective generator). *There exists an operator \mathcal{L}_{eff} acting on a vector space of measurable test functions of E_{eff} denoted \mathcal{D}_{eff} , and such that:*

i) \mathcal{D}_{eff} is a subset of continuous and bounded functions, contains constants, and contains a countable subset $(\varphi_i)_{i \in I}$ that separates points of E_{eff} .

ii) To any $\varphi \in \mathcal{D}_{\text{eff}}$, one can associate a bounded measurable function $\mathcal{L}_0 \mathcal{P} \varphi : E \rightarrow \mathbb{R}$ such that for any $\gamma > 0$ and any initial condition $x \in E$,

$$t \mapsto \mathcal{P} \varphi(X_t^{(\gamma,1)}(x)) - \int_0^t \mathcal{L}_0 \mathcal{P} \varphi(X_s^{(\gamma,1)}(x)) ds$$

is a martingale for the natural filtration of $X^{(\gamma,1)}(x)$ (Nota Bene: one has formally $(\gamma \mathcal{L}_1 + \mathcal{L}_0) \mathcal{P} = \mathcal{L}_0 \mathcal{P}$).

iii) For all $\varphi \in \mathcal{D}_{\text{eff}}$, $\mathcal{L}_0 \mathcal{P} \varphi$ is continuous in E at points of E_{eff} . Explicitly: for any $\varphi \in \mathcal{D}_{\text{eff}}$ and any $x_0 \in E_{\text{eff}}$, it holds $\mathcal{L}_{\text{eff}}[\varphi](x_0) = \lim_{x \rightarrow x_0} \mathcal{L}_0 \mathcal{P}[\varphi](x)$.

Remark 2.6. Point ii) in Assumption 7 is a rigorous way to express that $\mathcal{P} \varphi$ belongs to the domain of the generator $\gamma \mathcal{L}_1 + \mathcal{L}_0$ of the initial process. In order to check it, one may compute the predictable plus martingale decomposition (the so-called Doob-Meyer decomposition) of $t \mapsto \mathcal{P} \varphi(X_t^{(\gamma,1)})$ using e.g. Itô formula or an appropriate variant. Then one needs to identify the predictable part to the quantity $\int_0^t \mathcal{L}_0 \mathcal{P} \varphi(X_s^{(\gamma,1)}) ds$. Typically, $\mathcal{P} \varphi$ will belong to a 'natural' domain of the operator \mathcal{L}_0 (e.g. differentiable if \mathcal{L}_0 is a first order differential operator on a manifold). Note that defining $\mathcal{L}_0 \mathcal{P} \varphi$ outside of $E \setminus E_{\text{eff}}$ or even in a pointwise sense might not always be straightforward, and one may take the limit considered in Assumption 7 point iii) as a definition of $\mathcal{L}_0 \mathcal{P} \varphi$ on E_{eff} .

Finally, the following final assumption enables to identify the distribution of the effective Markov process on E_{eff} , using the classical characterization of Markov processes by martingale problems (see e.g. [11]). We recall that the distribution of a càdlàg random process $(X_t)_{t \geq 0}$ is said to be a solution of the *martingale problem* associated with the generator (with its domain) $(\mathcal{L}, \mathcal{D})$ if for all test function $\varphi \in \mathcal{D}$, the process

$$t \mapsto \varphi(X_t) - \int_0^t \mathcal{L} \varphi(X_s) ds$$

is a martingale for the natural filtration of the process $(\sigma(X_s, s \leq t))_{t \geq 0}$. The domain is chosen such that φ and $\mathcal{L} \varphi$ are measurable and $s \mapsto \mathcal{L} \varphi(X_s)$ is time integrable; in the present case under Assumption 7, φ will be bounded and continuous and $\mathcal{L} \varphi$ will be bounded. The martingale problem is said to be *well-posed* if there exists a unique process distribution solution to the martingale problem. The latter then satisfies the Markov property, and the pair $(\mathcal{L}, \mathcal{D})$ can be interpreted as the infinitesimal generator of the process.

Assumption 8 (Well-posedness of the effective martingale problem). *Let $\mu_0 = \lim_{\gamma \rightarrow +\infty} \mu_0^\gamma$ be given in $\text{Pr}(E)$. The martingale problem $(\mathcal{L}_{\text{eff}}, \mathcal{D}_{\text{eff}})$ for càdlàg processes with initial condition $\mu_0 \mathcal{P} \in \text{Pr}(E_{\text{eff}})$ is well-posed.*

Finally, note that it may be possible that in some cases, slightly different types of assumptions may be used:

Remark 2.7 (On the weakening of assumptions). It may happen that in some specific cases another method enables to obtain the uniform convergence result of Proposition 2.4 (a situation we call case i) in this remark). Even more directly, it may happen in some specific cases that another method enables to prove Step 2 in the proof of the main homogenization theorem. Step 2 mainly states that the effective process must lie in E_{eff} (a situation we call we call case ii) in this remark). We do not consider such examples in the present work, but several assumptions may then be weakened, and our proof can then be revisited with less constraints.

- If case i) (or ii) happens, Assumption 1 can be weakened to convergence in distribution only. Assumption 2 can also be weakened to cases where \mathcal{P} is only continuous at points of E_{eff} . The proof is similar but one has to restrict in Theorem 2.10 to initial conditions μ_0^γ that are either constant (independent of γ) or that have a limit when $\gamma \rightarrow +\infty$ whose support lies in E_{eff} .
- If case i) (or ii) happens, Assumption 4 and Assumption 5 are no longer required.
- If case ii) happens, Assumption 6 is also no longer necessary. The only three remaining important points to be checked are: first the containment in compact

sets (Assumption 3), second the continuity of both $\mathcal{P}\varphi$ and $\mathcal{L}_0\mathcal{P}\varphi$ at points of E_{eff} , and third the well-posedness of the effective martingale problem.

2.3 The case of diffusions with Lipschitz coefficients

Let us now discuss the assumptions of the previous section for a specific class of processes. In Section 4, we will present three specific examples that all belong to this class. All assumptions will be verified in detail for those examples; the goal of the present section being devoted to informal comments aimed at making the assumptions more transparent.

The process $t \mapsto X_t^{(\gamma_1, \gamma_0)}(x)$ is assumed to be a diffusion defined on a closed domain $E \subset \mathbb{R}^n$ as the strong solution of a standard Stochastic Differential Equation (SDE) of the general form:

$$\begin{aligned} dX_t^{(\gamma_1, \gamma_0)} &= \sqrt{\gamma_1}\sigma_1(X_t^{(\gamma_1, \gamma_0)})dW_t^{(1)} + \gamma_1 b_1(X_t^{(\gamma_1, \gamma_0)})dt \\ &\quad + \sqrt{\gamma_0}\sigma_0(X_t^{(\gamma_1, \gamma_0)})dW_t^{(0)} + \gamma_0 b_0(X_t^{(\gamma_1, \gamma_0)})dt, \end{aligned}$$

in which $t \mapsto (W_t^{(0)}, W_t^{(1)})$ are two \mathbb{R}^m -valued independent standard Brownian motions for some $m \geq 1$ and $(\sigma_1, b_1, \sigma_0, b_0)$ are bounded, globally Lipschitz (for the usual Euclidean metric) coefficients defined at least on E . Note that it is necessary at this stage to check that solutions of the considered SDE indeed remain in E for all time. In order to fix ideas, we give in Lemma B.1 (in appendix) an example of a simple general argument that enables to check this fact. We also refer to Section 4 for examples.

The first thing to check is that the dominant process almost surely converges in large time to a random point $X_\infty^{(1,0)}(x)$ of a closed subset $E_{\text{eff}} \subset E$ (the effective space) which is typically a smooth sub-manifold. One also needs to check that this random point is precisely x if x lies in the effective space E_{eff} (Assumption 1). This is typically associated with the fact that the coefficients (σ_1, b_1) defining the dominant process both vanish on E_{eff} and only on E_{eff} . To see this, it may be useful to perform an appropriate random time change (singular at E_{eff}) so that the time-changed dominant process almost surely hits E_{eff} in finite time for any initial condition.

Generally, the definition of the (possibly time-changed) dominant process gives easily that $X_\infty^{(1,0)}(x)$ is not deterministic unless the initial condition x lies in E_{eff} (Assumption 5). Regarding the continuity of the distribution of $X_\infty^{(1,0)}(x)$ with respect to the initial condition x (Assumption 2), it may rather be checked with Assumption 7 discussed below.

The containment in a compact set with high probability (Assumption 3) may be proven by resorting to a standard Lyapounov function estimate which must be uniform with respect to γ_1 . It is automatically satisfied if E is compact, although we stress our setting is not restricted to this case (see the second example in Section 4).

Next, Assumption 4 and Assumption 6 are direct consequences of the fact that $X^{(\gamma_1, \gamma_0)}$ is a strong solution of a SDE with Lipschitz coefficients (see Lemma B.2 in appendix).

We can now focus on Assumption 7. The associated space of test functions is usually the space of compactly supported smooth functions $\mathcal{D}_{\text{eff}} = C_c^\infty(E_{\text{eff}})$, which has a countable (point) separating subset given in a local chart by polynomials with rational coefficients (point i) of Assumption 7). In order to prove ii) in Assumption 7 (and as a matter of fact to moreover check Assumption 2), it is then sufficient to check that $x \mapsto \mathcal{P}\varphi(x)$ is, say, twice differentiable so that one can apply Itô formula. In some cases, one way to prove the regularity of $\mathcal{P}\varphi$ is to resort to Partial Differential Equations (PDE) regularity theory and identify $\mathcal{P}\varphi$ as the unique solution on E of the (e.g. (hypo)elliptic) Partial Differential Equation $a \times \mathcal{L}_1\psi = 0$ with smooth Dirichlet boundary conditions

$\psi = \varphi$ at E_{eff} and a well-chosen time-change scalar field $a : E \setminus E_{\text{eff}} \rightarrow]0, +\infty[$ (singular near E_{eff} (we recall that $\mathcal{P}\varphi$ is a solution to $\mathcal{L}_1\mathcal{P}\varphi = 0$ in an appropriate sense). More direct methods may be possible (See Section 4). The (Doob-Meyer) finite variation part of $t \mapsto \mathcal{P}\varphi(X_t^{(\gamma,1)})$ is eventually given by $\int_0^t \mathcal{L}_0\mathcal{P}\varphi(X_s^{(\gamma,1)})ds$. The regularity of $\mathcal{P}\varphi$ furthermore gives us that $\mathcal{L}_0\mathcal{P}\varphi$ is continuous on E (point iii) of Assumption 7). The value of $\mathcal{L}_0\mathcal{P}\varphi$ on E_{eff} then defines the effective generator $\mathcal{L}_{\text{eff}}[\varphi]$. In the present context, $\mathcal{P}\varphi$ is the solution of a (hypo)elliptic Dirichlet problem, and $\mathcal{L}_0\mathcal{P}\varphi$ may be interpreted as a "generalized Dirichlet-to-Neumann" operator. In the case where σ_0 vanishes on E_{eff} so that \mathcal{L}_0 is a first order differential operator, $\mathcal{L}_0\mathcal{P}\varphi$ is typically the generator of a Levy-type jump process.

Finally, one needs to prove the well-posedness of the martingale problem associated with $\mathcal{L}_0\mathcal{P}$ on E_{eff} . If one can describe the effective generator as associated with a strong solution of a stochastic differential equation with Lipschitz coefficients and driven by, say, a Levy homogeneous process, some quite generic results can give the well-posedness of the martingale problem (see e.g. the reference [17]).

2.4 Main homogenization theorem

In order to state the main theorem, we first define the pseudo-path topology on path space used in strong noise homogenization problems.

Definition 2.8 (Pseudo-paths). *Let E denotes a Polish space. To each measurable path $x : \mathbb{R}_+ \rightarrow E$, one can associate a probability on $E \times \mathbb{R}_+$ defined by:*

$$\delta_{x_t}(dx)e^{-t}dt \in \text{Pr}(E \times \mathbb{R}_+). \tag{2.4}$$

The subset of probability distributions in $\text{Pr}(E \times \mathbb{R}_+)$ of the form (2.4) is called the pseudo-paths space. It is closed for the usual (Polish) topology of $\text{Pr}(E \times \mathbb{R}_+)$ given by convergence in distribution. The pull-back on the equivalence classes of measurable paths equal Lebesgue almost everywhere will be called the pseudo-path topology. The associated space will be denoted $L^0(e^{-t}dt, E)$.

Remark 2.9 (Basic properties). The fact that the space of pseudo-paths is closed for convergence in distribution and other main results related to this topology are summarized in Section A. In particular, let us recall that pseudo-paths topology inherits the usual sequential characterization of convergence in distribution: a sequence of paths $(x^n)_n$ converges for the pseudo-paths topology to a path x^∞ if and only if $\lim_n \int \varphi(x_t^n, t)e^{-t}dt = \int \varphi(x_t^\infty, t)e^{-t}dt$ for all continuous and bounded φ . This convergence is in fact equivalent to the (a priori) stronger convergence "in probability" (also called "in measure") defined by $\lim_n \int_0^\infty \min(1, d(x_t^n, x_t^\infty))e^{-t}dt$ where d metrizes E . For that reason, the pseudo path space is denoted $L^0(e^{-t}dt, E)$.

We can now state our main theorem:

Theorem 2.10 (Strong noise homogenization). *Consider \mathcal{L}_1 and \mathcal{L}_0 two Markov generators, and $(\mu_0^\gamma)_{\gamma>0}$ a family of initial distributions with support in E converging in distribution towards μ_0 a distribution on E . Let Assumptions 1-2-3-4-5-6-7-8 be fulfilled. Then a càdlàg process $X^{(\gamma,1)}$ on E solution of the martingale problem of generator $\gamma\mathcal{L}_1 + \mathcal{L}_0$ with initial condition μ_0^γ converges in distribution when $\gamma \rightarrow +\infty$ for the (Meyer-Zheng) pseudo-path topology to the unique càdlàg Markov process X^∞ on E_{eff} solution of the martingale problem of generator \mathcal{L}_{eff} and initial condition $\mu_0\mathcal{P}$.*

The convergence in law of X^γ towards X^∞ for the pseudo-path topology implies by general considerations of the pseudo-path topology the convergence in law of $(X_{t_1}^\gamma, \dots, X_{t_k}^\gamma) \in E^k$ for Lebesgue almost all time sequences (t_1, \dots, t_k) in \mathbb{R}_+^k . We can in fact prove that the convergence is actually true for all time sequences: finite-dimensional distributions do converge.

Corollary 2.11 (Convergence of finite-dimensional distributions). *Consider the same notations as in Theorem 2.10. Under the same assumptions, and for all $(t_1, \dots, t_k) \in \mathbb{R}_+^k$, $k \geq 1$, the random sequence $(X_{t_1}^{(\gamma,1)}, \dots, X_{t_k}^{(\gamma,1)})$ converges in law to $(X_{t_1}^\infty, \dots, X_{t_k}^\infty)$.*

3 Proof of the main theorems

In order to prove Theorem 2.10, we will proceed in four steps. First, using a tightness argument, we will extract a sub-sequence $(X^{\gamma_n})_{n \geq 1}$ of the considered family of initial processes which is converging in distribution for the pseudo-path topology. We denote the extracted limit X^∞ , and we underline that X^∞ is really defined as a pseudo-path, that is as the random probability distribution $\delta_{X_t^\infty}(dx)e^{-t}dt$. Next, we will prove that X^∞ takes its values in E_{eff} , which amounts to say that almost surely, $X_t^\infty \in E_{\text{eff}}$ for Lebesgue almost all $t > 0$. Then, using the so-called Meyer-Zheng criterion (see Section A), we will prove that X^∞ can be identified with a càdlàg process, in the sense that the distribution of X^∞ is the push-forward in pseudo-path space of the distribution of a càdlàg random process. Eventually, we will prove that X^∞ is a solution of a martingale problem on E_{eff} (which have a unique solution by assumption). We conclude that $X^{(\gamma,1)}$ converges in law to X^∞ for the pseudo-path topology when γ goes to infinity. We stress that the third step is unavoidable: limits with respect to the pseudo-path topology, *without the Meyer-Zheng criterion*, may be too degenerate to handle the convergence of the martingale part and obtain that the extracted limit indeed satisfies the expected martingale problem.

In the penultimate section we will prove that we have furthermore convergence of all finite-dimensional distributions.

In the last section, we will prove that under our assumptions the large time convergence of the dominant process is uniform with respect to initial conditions taken in a compact set, as stated in Proposition 2.4; a result used in Step 2.

3.1 Step 1: Tightness and extraction

We start by proving that, under Assumption 3, the considered family of processes is tight when considering the pseudo-path topology.

Lemma 3.1. *Under Assumption 3, the family $(X^{(\gamma,1)})_{\gamma > 0}$ with converging initial distribution is tight under the pseudo-path topology.*

Proof. We start by proving the following claim: let $(K_n)_{n \geq 0}$ denotes any given increasing sequence of compact subsets of E , then the set of pseudo-paths defined by:

$$\mathcal{K} := \bigcap_{n \geq 0} \left\{ x : \int_0^n \mathbb{1}_{x_t \in K_n^c} e^{-t} dt = 0 \right\}, \tag{3.1}$$

that is the set of pseudo-paths that remains in K_n on the time interval $[0, n]$ for all n , is relatively compact for the pseudo-path topology. In order to prove this claim, let us consider \mathcal{K} as a subset of the space $\text{Pr}(E \times \mathbb{R}_+)$, that is we identify $x \in \mathcal{K}$ with the probability $\delta_{x_t}(dy)e^{-t}dt$. Let $\varepsilon > 0$ be given, arbitrarily small. Let n_ε be such that $e^{-n_\varepsilon} \leq \varepsilon$. One has, by construction of \mathcal{K} , that for any $x \in \mathcal{K}$

$$\int_E \int_0^\infty \mathbb{1}_{(y,t) \in (K_{n_\varepsilon} \times [0, n_\varepsilon])^c} \delta_{x_t}(dy) e^{-t} dt \leq e^{-n_\varepsilon} \leq \varepsilon.$$

By Prokhorov theorem, \mathcal{K} is thus relatively compact, and the initial claim is proved.

We can now proceed and prove that under Assumption 3, the considered family of processes is tight under the pseudo-path topology.

Let $\varepsilon > 0$ be given, arbitrarily small. By the claim proved above, it is sufficient in order to conclude, to construct a n -increasing sequence $K_{\varepsilon,n} \subset E$ of compact sets such that

$$\sup_{\gamma} \mathbb{P} \left[X^{(\gamma,1)} \in \mathcal{K}_{\varepsilon}^c \right] \leq \varepsilon,$$

where the relatively compact set $\mathcal{K}_{\varepsilon}$ is defined using the sequence $(K_{\varepsilon,n})_{n \geq 1}$ and (3.1). For this purpose, using Assumption 3 with variant (2.3), we choose $K_{\varepsilon,n}$ such that we have $\mathbb{P}(\exists t \in [0, n], X_t^{(\gamma,1)} \in K_{\varepsilon,n}^c) \leq \varepsilon 2^{-n}$ uniformly in γ .

Then by construction:

$$\mathbb{P} \left[X^{(\gamma,1)} \in \mathcal{K}_{\varepsilon}^c \right] \leq \sum_n \mathbb{P} \left[\exists t \in [0, n], X_t^{(\gamma,1)} \in K_{\varepsilon,n}^c \right] \leq \varepsilon.$$

□

As a consequence, any increasing sequence converging to infinity has an increasing sub-sequence denoted $(\gamma_n)_{n \geq 0}$ such that $X^{(\gamma_n,1)}$, when embedded as a pseudo-path in $\text{Pr}(E \times \mathbb{R}_+)$, converges when n goes to infinity in distribution towards a limit. Since we know that the set of pseudo-paths is closed, this limit is the distribution of a pseudo-path and can be represented as a random process denoted X^∞ . However, X^∞ may not have a càdlàg representative at this stage (this will be taken care of in Step 3 below).

3.2 Step 2: The limit process lies in the effective space

The goal of this step is to prove that the considered limit process takes its values in E_{eff} in the sense that:

$$\mathbb{E} \left[\int_0^{+\infty} \mathbb{1}_{X_t^\infty \in E_{\text{eff}}} e^{-t} dt \right] = 1.$$

We use Proposition 2.4 to prove that the full process approaches E_{eff} when γ approaches infinity:

Lemma 3.2. *Let f denotes the bounded continuous non-negative function on E vanishing on E_{eff} and strictly positive on $E \setminus E_{\text{eff}}$ characterizing the uniform convergence of the dominant process towards E_{eff} in Proposition 2.4. Let μ_0^γ denotes the initial condition of $X_t^{(\gamma,1)}$. Under Assumptions 3 and 6, we have for all $t > 0$ convergence towards E_{eff} for large γ :*

$$\mathbb{E}_{\mu_0^\gamma}(f(X_t^{(\gamma,1)})) \xrightarrow{\gamma \rightarrow +\infty} 0.$$

Proof. First notice that

$$\begin{aligned} \mathbb{E}_{\mu_0^\gamma}(f(X_t^{(\gamma,1)})) &= \langle \mu_0^\gamma, e^{t(\gamma\mathcal{L}_1 + \mathcal{L}_0)} f \rangle \\ &= \langle \mu_0^\gamma, e^{\gamma t(\mathcal{L}_1 + \frac{1}{\gamma}\mathcal{L}_0)} f \rangle \\ &= \mathbb{E}_{\mu_0^\gamma}(f(X_{\gamma t}^{(1, \frac{1}{\gamma})})), \end{aligned}$$

which amounts to an elementary time change.

Therefore, for any $0 < \delta < \gamma t$ we have:

$$\begin{aligned} & \left| \mathbb{E}_{\mu_0^\gamma} \left(f \left(X_{\gamma t}^{(1, \frac{1}{\gamma})} \right) \right) \right| \\ & \leq \underbrace{\left| \mathbb{E}_{\mu_0^\gamma} \left(f \left(X_{\gamma t}^{(1, \frac{1}{\gamma})} \right) \right) - \mathbb{E}_{\mu_0^\gamma} \left(f \left(X_{\gamma t}^{(1,0)} \right) \right) \right|}_{= \left| \mathbb{E}_{\mu_0^\gamma} \left[\mathbb{E}_{X_{\gamma t - \delta}^{(1, \frac{1}{\gamma})}} \left(f \left(X_{\delta}^{(1, \frac{1}{\gamma})} \right) \right) - \mathbb{E}_{X_{\gamma t - \delta}^{(1,0)}} \left(f \left(X_{\delta}^{(1,0)} \right) \right) \right]} \right| + \left| \mathbb{E}_{\mu_0^\gamma} \left(f \left(X_{\gamma t}^{(1,0)} \right) \right) \right| \end{aligned}$$

As a consequence, for any compact subset K of E

$$\begin{aligned}
 & \left| \mathbb{E}_{\mu_0^\gamma} \left(f \left(X_{\gamma t}^{(1, \frac{1}{\gamma})} \right) \right) \right| \\
 & \leq \left| \mathbb{E}_{\mu_0^\gamma} \left[\mathbb{E}_{X_{\gamma t - \delta}^{(1, \frac{1}{\gamma})}} \left(f \left(X_\delta^{(1, \frac{1}{\gamma})} \right) \right) - \mathbb{E}_{X_{\gamma t - \delta}^{(1, \frac{1}{\gamma})}} \left(f \left(X_\delta^{(1, 0)} \right) \right) \right] \right| \\
 & \quad + \left| \mathbb{E}_{\mu_0^\gamma} \left[\mathbb{E}_{X_{\gamma t - \delta}^{(1, \frac{1}{\gamma})}} \left(f \left(X_\delta^{(1, 0)} \right) \right) - \mathbb{E}_{X_{\gamma t - \delta}^{(1, 0)}} \left(f \left(X_\delta^{(1, 0)} \right) \right) \right] \right| \\
 & \quad + \left| \mathbb{E}_{\mu_0^\gamma} \left(f \left(X_{\gamma t}^{(1, 0)} \right) \right) \right| \\
 & \leq \sup_{x \in K} \left| \mathbb{E} \left(f \left(X_\delta^{(1, \frac{1}{\gamma})} \right) (x) \right) - \mathbb{E} \left(f \left(X_\delta^{(1, 0)} \right) (x) \right) \right| + 2 \|f\|_\infty \mathbb{P}_{\mu_0^\gamma} \left[X_{t - \delta/\gamma}^{(\gamma, 1)} \in K^c \right] \\
 & \quad + 2 \times \sup_{x \in K} \mathbb{E} \left[f \left(X_\delta^{(1, 0)} \right) (x) \right] + \|f\|_\infty \mathbb{P}_{\mu_0^\gamma} \left[X_{t - \delta/\gamma}^{(\gamma, 1)} \in K^c \right] + \|f\|_\infty \mathbb{P}_{\mu_0^\gamma} \left[X_{t - \delta/\gamma}^{(\gamma, 0)} \in K^c \right] \\
 & \quad + \sup_{x \in K} \left| \mathbb{E} \left(f \left(X_{\gamma t}^{(1, 0)} \right) (x) \right) \right| + \|f\|_\infty \mu_0^\gamma(K^c) \tag{3.2}
 \end{aligned}$$

Let $\varepsilon > 0$ be given, arbitrarily small. Using Assumption 3 (compact containment), one can choose $K = K_\varepsilon$ such that, uniformly in δ, γ :

$$\max \left(\mathbb{P}_{\mu_0^\gamma} \left[X_{t - \delta/\gamma}^{(\gamma, 1)} \in K_\varepsilon^c \right], \mathbb{P}_{\mu_0^\gamma} \left[X_{t - \delta/\gamma}^{(\gamma, 0)} \in K_\varepsilon^c \right], \mu_0^\gamma(K_\varepsilon^c) \right) \leq \frac{\varepsilon}{10 \|f\|_\infty}.$$

Using Proposition 2.4 (uniform time convergence of the dominant process), we consider now $\delta_\varepsilon = \delta_{\varepsilon, K_\varepsilon}$ large enough such that

$$\sup_{x \in K_\varepsilon} \mathbb{E} \left[f \left(X_{\delta_\varepsilon}^{(1, 0)} \right) (x) \right] \leq \frac{\varepsilon}{10}$$

Now using Assumption 6 (uniformly continuous perturbation), we now consider $\gamma_\varepsilon := \gamma_{\varepsilon, K_\varepsilon} > 0$ such that for all $\gamma > \gamma_\varepsilon$:

$$\sup_{x \in K_\varepsilon} \left| \mathbb{E} \left(f \left(X_{\delta_\varepsilon}^{(1, 1/\gamma)} \right) (x) \right) - \mathbb{E} \left(f \left(X_{\delta_\varepsilon}^{(1, 0)} \right) (x) \right) \right| \leq \frac{\varepsilon}{5}$$

Taking $\delta = \delta_\varepsilon$ in inequality (3.2) thus implies for all $\gamma > \max \left(\frac{\delta_\varepsilon}{t}, \gamma_\varepsilon \right)$:

$$\left| \mathbb{E}_{\mu_0^\gamma} \left(f \left(X_t^{(\gamma, 1)} \right) \right) \right| \leq \varepsilon.$$

□

We have proved in Step 1 that there exists $(\gamma_n)_n$ an increasing sequence of positive numbers approaching infinity such that we have $X^{(\gamma_n, 1)} \xrightarrow{\mathcal{L}} X^\infty$ for the pseudo-path topology. By characterization of the pseudo-path convergence, it implies that

$$\mathbb{E} \int_0^{+\infty} f \left(X_t^{(\gamma_n, 1)} \right) e^{-t} dt \xrightarrow{n \rightarrow +\infty} \mathbb{E} \int_0^{+\infty} f \left(X_t^\infty \right) e^{-t} dt.$$

By dominated convergence and Lemma 3.2 above, the term on the left converges to 0. But $f^{-1}(\{0\}) = E_{\text{eff}}$ and $f \geq 0$, so eventually $X^\infty \in E_{\text{eff}}$ almost surely as a pseudo-path (that is for Lebesgue almost all t).

3.3 Step 3: The limit is càdlàg

We will now prove that the distribution of X^∞ has a càdlàg representative taking values in E_{eff} . To do so, we will use the so-called *Meyer-Zheng criterion*. The latter ensures that a sequence of real valued càdlàg random processes that i) converges in distribution for the pseudo-path topology, and ii) has a uniformly bounded *mean variation* (see below for a definition), has a pseudo-path limit which possesses a càdlàg representative.

Definition 3.3 (Mean variation). *Let $T > 0$ denotes an horizon time, and $(Z_t)_{t \geq 0}$ a càdlàg real valued random process with $Z_t \in L^1(\mathbb{P})$ for all t . The mean variation of Z over $[0, T]$ is given by*

$$V_T(Z) := \sup_{0=t_0 \leq t_1 \dots \leq t_K \leq T} \sum_{k=0}^K \mathbb{E} \left| \mathbb{E} [Z_{t_{k+1}} - Z_{t_k} \mid \sigma(Z_t, t \leq t_k)] \right|,$$

$$= \sup_{|H| \leq 1} \mathbb{E} \left(\int_0^T H_t dZ_t \right),$$

where in the above the supremum is taken over predictable processes (with respect to the natural filtration of Z) taking value in $[-1, 1]$.

We recall in Section A several general facts underlying the Meyer-Zheng approach to limit theorems for processes. In particular, processes with bounded mean variations are quasi-martingales, and their bounded mean variation is precisely equal to the total variation of their predictable finite-variation part (defined in the sense of the Doob-Meyer decomposition of quasi-martingales).

Step 3 (the limit process X^∞ has a – unique – càdlàg representative) will be a direct consequence of Lemma 3.5 below. We start with a technical remark that will ensure generic measurability of some random objects used in the proof of Lemma 3.5.

Lemma 3.4. *Let I be a countable set and \mathbb{R}^I be endowed with the usual product (Polish) topology. Let F denotes a Polish space, and let $\Psi : F \rightarrow \mathbb{R}^I$, be an injective bounded continuous function. Consider the weaker topology on F induced by the pull-back by Ψ , for instance as defined by the distance*

$$\tilde{d}(x, y) = \sum_{i \in I} 2^{-n_i} |\Psi_i(x) - \Psi_i(y)|, \tag{3.3}$$

where $i \mapsto n_i \in \mathbb{N}$ is injective. Denote by \tilde{F} the completion of F with respect to the latter distance. Then i) Ψ defines a homeomorphism on \tilde{F} for the weaker topology defined by \tilde{d} , and ii) all Borel sets of F defined by the strong original topology (defined by d) are also Borel in F for the weaker new topology (defined by \tilde{d}).

Proof. Point i) is trivial: \tilde{d} is a metric because the family $(\Psi_i)_{i \in I}$ is separating. Then \tilde{F} is simply identified with $\overline{\Psi(F)}$ by the isometry defined by \tilde{d} .

Point ii) seems obvious but is more subtle. A standard result in descriptive set theory ([16, Theorem 15.1]) states that on Polish spaces, injective Borel maps are in fact Borel isomorphisms, in the sense that they map Borel sets to Borel sets. We know that $\Psi : (F, d) \rightarrow \mathbb{R}^I$ is injective and continuous, it is thus Borel and so is its inverse Ψ^{-1} defined on Polish $\Psi(F)$. Let $B \subset F$ be Borel for the strong topology $\Phi(B)$ is thus Borel in $\Psi(F)$; by isometry $B \subset F$ is eventually Borel for the topology induced by \tilde{d} . \square

Lemma 3.5. *Let X^n be a sequence of càdlàg processes on E converging in distribution to a pseudo-path X^∞ for the pseudo-path topology, and let $E_{\text{eff}} \subset E$. We assume the following:*

1. The sequence X^n satisfies the compact containment condition (2.3) (the index n playing the role of γ).
2. The limit pseudo-path X^∞ almost surely has support in E_{eff} .
3. There exists a countable family of bounded continuous map $\Psi_i : E \rightarrow \mathbb{R}$, $i \in I$ such that for all $T \geq 0$ and all $i \in I$:

$$\sup_n V_T(\Psi_i(X^n)) < +\infty.$$

In other words the (quasi-martingale) càdlàg processes $\Psi_i(X^n)$ have a uniformly (w.r.t. n) bounded mean variation on finite time intervals (Meyer–Zheng criterion).

4. The family $(\Psi_i)_{i \in I}$ separates the points of E_{eff} , that is $\Psi : E_{\text{eff}} \rightarrow \mathbb{R}^I$ is injective.

Then the distribution of X^∞ is given by the distribution of a càdlàg process taking its values in E_{eff} .

Proof. For each $i \in I$, by continuity of Ψ_i , the distribution of the bounded random pseudo-path $\Psi_i(X^n) \in \mathbb{R}$ converges when $n \rightarrow +\infty$ towards the distribution of the random pseudo-path $\Psi_i(X^\infty)$. Point 3 in the assumption is then exactly the uniform mean variation condition of Meyer-Zheng theory (see Theorem A.6 in Section A), which ensures that the distribution of the pseudo-path $\Psi_i(X^\infty)$ is in fact supported by $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$ (the space of càdlàg paths on \mathbb{R}). This implies existence of a unique càdlàg representative of $\Psi_i(X^\infty)$, given and denoted (to avoid confusion) by

$$Z_t^i := \lim_{h \rightarrow 0^+} \frac{1}{h} \int_t^{t+h} \Psi_i(X_s^\infty) ds.$$

Note that if one considers a measurable process $(t, \omega) \mapsto X_t^\infty(\omega) \in E_{\text{eff}}$ taking values in E_{eff} and representing X^∞ (see Lemma A.2 for existence of such representative), then the set of elements $(\omega, t) \in \Omega \times \mathbb{R}_+$ such that $\Psi_i(X_t^\infty(\omega)) = Z_t^i(\omega)$ has $e^{-t} dt \otimes \mathbb{P}$ -measure 1, and since I is countable, the intersection for $i \in I$ is again of $e^{-t} dt \otimes \mathbb{P}$ -measure 1.

Using Point 4 (Ψ is injective on E_{eff}) and Lemma 3.4, one can then consider the extension of the inverse $\Psi^{-1} : \Psi(E_{\text{eff}}) \rightarrow \tilde{E}_{\text{eff}}$ which defines an isometry when one considers e.g. the distance (3.3) associated with the product Polish topology of \mathbb{R}^I . With respect to this weaker topology, the process

$$t \mapsto \tilde{X}_t^\infty := \Psi^{-1}(Z_t) \in \tilde{E}_{\text{eff}} \supset E_{\text{eff}}$$

is almost surely càdlàg.

Of course, \tilde{X}^∞ is a representative of the random pseudo-path X^∞ when considered as a pseudo-path taking values in E_{eff} (they are equal on a set of full $e^{-t} dt \otimes \mathbb{P}$ measure). In order to check this point rigorously while being careful with measurability, it is useful to note that by Lemma 3.4 point ii), Borel sets of E_{eff} for the initial stronger topology are also Borel for the induced weaker topology, so that if φ is a bounded function on E_{eff} Borel measurable for the strong topology, the integral $\int_0^{+\infty} \varphi(\tilde{X}_t, t) e^{-t} dt$ is well-defined simply by the fact the process $\tilde{X}_t \in \tilde{E}_{\text{eff}}$ lies in $E_{\text{eff}} \subset \tilde{E}_{\text{eff}}$ almost surely and for Lebesgue almost all t . This happens in the present case by construction so that one can write $\int_0^{+\infty} \varphi(\tilde{X}_t^\infty, t) e^{-t} dt = \int_0^{+\infty} \varphi(X_t^\infty, t) e^{-t} dt$, which rigorously justifies $\tilde{X}^\infty = X^\infty$ as pseudo-paths on E_{eff} endowed with the initial strong topology.

It remains, in order to conclude the proof, to show that \tilde{X}^∞ is càdlàg for the stronger original topology of E . Note that in the simpler case where E_{eff} is compact, the two topologies are equivalent $E_{\text{eff}} = \tilde{E}_{\text{eff}}$, and the proof is complete. In order to finish the

proof when E is not compact, we are to use a compactness argument (Point 1) and show that \tilde{X}^∞ has, almost surely, left and right accumulation points in the original stronger topology of E . This is not guaranteed in general without the compact containment condition because the weaker topology of \tilde{E}_{eff} may allow values in $\tilde{E}_{\text{eff}} \setminus E_{\text{eff}}$.

More precisely, we now claim that with the compact containment condition (Point 1), for any $T > 0$, the set $\{\tilde{X}_t^\infty(\omega), t \in [0, T]\}$ is $\mathbb{P}(d\omega)$ -almost surely a subset of E that is relatively compact for the stronger original topology. If this claim is true, $\mathbb{P}(d\omega)$ -almost surely, for any (left or right) converging sequence $t_m \rightarrow t_\infty$, there is a sub-sequence (we do not change notation for the sub-sequence) such that $\tilde{X}_{t_m}^\infty(\omega)$ converges for the stronger topology in E and lies in fact in E_{eff} since the set is closed. But $t \mapsto \tilde{X}_t^\infty(\omega)$ is càdlàg for the weaker topology of \tilde{E}_{eff} , so the limit is unique and given by $\tilde{X}_{t_\infty}^\infty(\omega)$. Hence \tilde{X}^∞ is càdlàg in E_{eff} (for the original strong topology) and the proof of the whole lemma is complete.

It remains to prove the claim above. We consider the original topology on E_{eff} . We first remark that (by portmanteau lemma) the map on pseudo-paths $x \mapsto \int_0^T \mathbb{1}_{x_t \in K^c} e^{-t} dt$ is lower semi-continuous for K closed, so that $\{x : \int_0^T \mathbb{1}_{x_t \in K^c} e^{-t} dt > 0\}$ is open and applying the portmanteau lemma in pseudo-path space now yields

$$\mathbb{P} \left[\int_0^T \mathbb{1}_{X_t^\infty \in K^c} e^{-t} dt > 0 \right] \leq \liminf_n \mathbb{P} \left[\int_0^T \mathbb{1}_{X_t^n \in K^c} e^{-t} dt > 0 \right].$$

The compact containment condition implies that the right-hand side in the above is smaller than any ε for a well chosen $K = K_\varepsilon$. Using a routine Borel-Cantelli argument obtained by considering the events constructed for $K_{2^{-p}}, p \in \mathbb{N}$, one gets that:

$$\mathbb{P} \left[\exists p \geq 0 : \int_0^T \mathbb{1}_{X_t^\infty \in K_{2^{-p}}^c} e^{-t} dt = 0 \right] = 1.$$

Since X^∞ is a pseudo-path of E_{eff} , we can here replace $K_{2^{-p}}$ by $K_{2^{-p}} \cap E_{\text{eff}}$ that is still a compact since E_{eff} closed. The above statement holds for X^∞ taking value in pseudo-paths space, so that one can replace X^∞ with \tilde{X}^∞ . In the weaker topology of \tilde{E}_{eff} , the compact $K_{2^{-k}} \cap E_{\text{eff}} \subset \tilde{E}_{\text{eff}}$ is again compact. Since \tilde{X}^∞ is càdlàg, it implies that $\mathbb{P}(d\omega)$ -almost surely for a random integer $P(\omega)$, we have $\tilde{X}_t^\infty(\omega) \in K_{P(\omega)}$ for all $t \in [0, T]$. This precisely means that the set $\{\tilde{X}_t^\infty(\omega), t \in [0, T]\}$ is relatively compact for the strong original topology of E_{eff} . Our claim is proved, and so is the lemma. □

Let us now apply Lemma 3.5 to our setting in order to prove Step 3.

Let φ_i denotes a countable family in \mathcal{D}_{eff} that separates points. We now consider

$$X^n = X^{(\gamma_n, 1)}, \quad \Psi_i = \mathcal{P}\varphi_i,$$

We have already proved in Step 2 the second condition of Lemma 3.5 while the first one is Assumption 3 and the fourth one is part i) of Assumption 7. For the third one (the most important), we write:

$$V_T(\Psi_i(X^n)) = \sup_{|H| \leq 1} \int_0^T \mathbb{E}(H_t d\Psi_i(X_t^n)),$$

where the supremum is taken over predictable processes (with respect to the natural filtration of X^n) taking value in $[-1, 1]$. Since the martingale term disappears by aver-

aging:

$$\begin{aligned} V_T(\Psi_i(X^n)) &= \sup_{|H| \leq 1} \int_0^T \mathbb{E}(H_t (\gamma_n \mathcal{L}_1 + \mathcal{L}_0) \mathcal{P}\varphi_i(X_t^n)) dt \\ &= \sup_{|H| \leq 1} \int_0^T \mathbb{E}(H_t \mathcal{L}_0 \mathcal{P}\varphi_i(X_t^n)) dt \\ &= \int_0^T \mathbb{E}(|\mathcal{L}_0 \mathcal{P}\varphi_i(X_t^n)|) dt \\ &\leq T \|\mathcal{L}_0 \mathcal{P}\varphi_i\|_\infty, \end{aligned}$$

this final quantity being finite by part ii) of Assumption 7. This yields the uniform mean variation condition. We stress that the above estimate is the *key technical ingredient of the strong homogenization theorem*.

Lemma 3.5 gives us therefore that X^∞ is almost surely càdlàg with trajectories in E_{eff} .

3.4 Step 4: The limit is the unique solution of a martingale problem

We can now prove that X^∞ is solution of a specific martingale problem. We already know by Assumption 7 that for all $\gamma > 0$ and $\varphi \in \mathcal{D}_{\text{eff}}$, $X^\gamma := X^{(\gamma, \cdot)}$ is solution of:

$$\mathcal{M}^\gamma : t \mapsto \mathcal{P}\varphi(X_t^\gamma) - \int_0^t \mathcal{L}_0 \mathcal{P}\varphi(X_s^\gamma) ds \text{ is a } \sigma(X_s^\gamma, s \leq t)\text{-martingale.}$$

We would like to let γ goes to infinity in this problem, and get that:

$$\mathcal{M}^\infty : t \mapsto \mathcal{P}\varphi(X_t^\infty) - \int_0^t \mathcal{L}_0 \mathcal{P}\varphi(X_s^\infty) ds \text{ is a } \sigma(X_s^\infty, s \leq t)\text{-martingale.}$$

To prove this, we will first prove using a general result (Lemma A.3) that Lebesgue almost all finite-dimensional distributions of \mathcal{M}^{γ_k} converge to those of \mathcal{M}^∞ . In particular this will enable to identify the distribution of the effective initial condition X_0^∞ . We will then invoke a routine lemma characterizing càdlàg martingales (Theorem A.7) in order to check that \mathcal{M}^∞ is indeed a martingale. The martingale problem uniqueness in Assumption 8 enables to identify the distribution of X^∞ and then to conclude.

Using the general Lemma A.3 for converging processes distributions with respect to the pseudo-path topology *towards a càdlàg supported limit*, we know that there exists a subset with full Lebesgue measure $J \subset \mathbb{R}_+$ and a sub-sequence $\gamma_n \rightarrow +\infty$ such that for each t_1, \dots, t_p the joint variable

$$(X^{\gamma_n}, X_{t_1}^{\gamma_n}, \dots, X_{t_p}^{\gamma_n}) \xrightarrow[n \rightarrow +\infty]{\text{Law}} (X^\infty, X_{t_1}^\infty, \dots, X_{t_p}^\infty)$$

converges in distribution in $L^0(e^{-t} dt, E) \times E^p$. Since by Assumption 7 point iii), $\mathcal{L}_0 \mathcal{P}\varphi$ is bounded and continuous in E at each point of E_{eff} , the map

$$x \mapsto \int_0^t \mathcal{L}_0 \mathcal{P}\varphi(x_s) ds$$

is bounded and continuous at points of $L^0(e^{-t} dt, E_{\text{eff}})$ for the pseudo-path topology. In the same way, Assumption 2 ensures that the map $x \mapsto \mathcal{P}\varphi(x)$ is bounded and continuous. Combining the above results, we get that Lebesgue almost all (for times taken in J) finite-dimensional distributions of \mathcal{M}^{γ_p} do converge to those of \mathcal{M}^∞ .

We now claim that this proves that:

$$\text{Law}(X_{0+}^\infty) = \mu_0 \mathcal{P}.$$

First recall that the trajectories of X^∞ are almost surely in E_{eff} , and we have almost surely that $\mathcal{P}\varphi(X^\infty) = \varphi(X^\infty)$. On the other hand by construction $\mathbb{E}[\mathcal{M}_t^\gamma] = \mathbb{E}(\mathcal{P}\varphi(X_0^\gamma)) = \mu_0^\gamma \mathcal{P}\varphi$ for each t , so that passing to the limit for $t \in \mathbb{J}$

$$\mathbb{E}[\mathcal{M}_t^\infty] = \mu_0 \mathcal{P}\varphi$$

by Assumption 2, and since $\mathcal{M}_{0+}^\infty = \varphi(X_{0+}^\infty)$ the claim is proved.

We now turn to the martingale property. Using the convergence of finite-dimensional distributions on J , one gets: for all $0 \leq t_1 < \dots < t_p < t$ in J , for all $\varphi_1, \dots, \varphi_p$ continuous bounded, and for all p :

$$\begin{array}{ccc} \mathbb{E}[(\mathcal{M}_t^\gamma - \mathcal{M}_{t_p}^\gamma)\varphi_p(X_{t_p}^\gamma)\dots\varphi_1(X_{t_1}^\gamma)] & = & 0 \\ \downarrow \gamma \rightarrow +\infty & & \downarrow \gamma \rightarrow +\infty \\ \mathbb{E}[(\mathcal{M}_t^\infty - \mathcal{M}_{t_p}^\infty)\varphi_p(X_{t_p}^\infty)\dots\varphi_1(X_{t_1}^\infty)] & = & 0. \end{array}$$

Since the process \mathcal{M}^∞ is a càdlàg process, we have \mathcal{M}^∞ is a martingale for the filtration of X^∞ by Theorem A.7.

Finally, recalling again that $\mathcal{P}\varphi(X^\infty) = \varphi(X^\infty)$ (trajectories of X^∞ are almost surely in E_{eff}), we have that:

$$t \mapsto \varphi(X_t^\infty) - \int_0^t \mathcal{L}_0 \mathcal{P}\varphi(X_s^\infty) \text{ is a } \sigma(X_s^\infty, s \leq t)\text{-martingale.}$$

Assumption 8 ensures eventually that the distribution of X^∞ is the unique càdlàg solution of the above martingale problem.

Routinely combining this identification of the limit with the tightness of $(X^\gamma)_{\gamma>0}$ in pseudo-path space obtained in Step 1, we find that X^γ indeed converges in law for the pseudo-path topology to X^∞ , the unique (in law) process on E_{eff} of generator \mathcal{L}_{eff} : the Theorem 2.10 is proved.

3.5 Proof of Corollary 2.11

In order to prove the convergence of any finite-dimensional marginal distribution of the process, we first need to prove the convergence of a single time marginal distribution starting from a given moving initial condition.

Proposition 3.6. *We assume as before that X^γ starts from the initial condition μ_0^γ that converges in law towards μ_0 . Then for all $t > 0$, under the assumptions of Theorem 2.10, we have that X_t^γ converges in law towards X_t^∞ of initial condition $\mu_0 \mathcal{P}$.*

Proof. Let $t > 0$ be given. The compact containment condition of Assumption 3 implies that the family of distributions $(X_t^\gamma)_\gamma$ is tight, and we can thus consider a first extraction converging in distribution towards a limit denoted Y_t . By Lemma 3.2, we know that this limit lies in the effective state space $Y_t \in E_{\text{eff}}$ almost surely.

Next, as in the proof of Step 4 above, since X^γ converges in law to càdlàg X^∞ for the pseudo-path topology, there exists a second extraction (γ_n) and $\mathbb{J} \subset \mathbb{R}_+$ of full Lebesgue measure such that the finite-dimensional distributions of $(X_s^{\gamma_n})_{s \in \mathbb{J}}$ converge to those of $(X_s^\infty)_{s \in \mathbb{J}}$.

Consider a sequence $(t_k)_k$ of elements of \mathbb{J} converging to t from the right. For $f \in \mathcal{D}_{\text{eff}}$ (by assumption a separating class of bounded test functions), we can now consider the following decomposition (we recall that $f = \mathcal{P}f$ on E_{eff}):

$$\begin{aligned} & |\mathbb{E}_{\mu_0^{\gamma_n}}(\mathcal{P}f(X_t^{\gamma_n})) - \mathbb{E}_{\mu_0 \mathcal{P}}(f(X_t^\infty))| \\ & \leq |\mathbb{E}_{\mu_0^{\gamma_n}}(\mathcal{P}f(X_t^{\gamma_n}) - \mathcal{P}f(X_{t_k}^{\gamma_n}))| + |\mathbb{E}_{\mu_0^{\gamma_n}}(\mathcal{P}f(X_{t_k}^{\gamma_n})) - \mathbb{E}_{\mu_0 \mathcal{P}}(\mathcal{P}f(X_{t_k}^\infty))| \\ & \quad + |\mathbb{E}_{\mu_0 \mathcal{P}}(\mathcal{P}f(X_{t_k}^\infty)) - \mathbb{E}_{\mu_0 \mathcal{P}}(f(X_t^\infty))|. \end{aligned} \tag{3.4}$$

First note that since $\mathcal{P}f$ is continuous (Assumption 2) and bounded the left hand side converges for n large towards $|\mathbb{E}(f(Y_t)) - \mathbb{E}_{\mu_0\mathcal{P}}(f(X_t^\infty))|$.

We also remind that almost surely

$$f(X_{t_k}^\infty) \xrightarrow[k \rightarrow +\infty]{} f(X_t^\infty)$$

by continuity of f and right continuity of $s \mapsto X_s^\infty$. Since f is bounded, we have by dominated convergence:

$$\mathbb{E}_{\mu_0\mathcal{P}}(f(X_{t_k}^\infty)) \xrightarrow[k \rightarrow +\infty]{} \mathbb{E}_{\mu_0\mathcal{P}}(f(X_t^\infty)).$$

Using this last result and point ii) of Assumption 7, we now consider k such that $|\mathbb{E}_{\mu_0\mathcal{P}}(f(X_{t_k}^\infty)) - \mathbb{E}_{\mu_0\mathcal{P}}(f(X_t^\infty))| \leq \varepsilon$ and $(t_k - t) \sup_E |\mathcal{L}_0\mathcal{P}f| \leq \varepsilon$. The third term of the right hand side of (3.4) is lower than ε uniformly in n . We fix $k := k_\varepsilon$ until the end of the proof. It remains (before concluding) to prove that the two other terms are inferior to ε for any n large enough.

Since $t_k \in \mathbb{J}$ we have $X_{t_k}^{\gamma_n} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X_{t_k}^\infty$ and we bound the second term by ε for n large enough.

For the first term, using Dynkin's formula (since $(\gamma\mathcal{L}_1 + \mathcal{L}_0)\mathcal{P}f = \mathcal{L}_0\mathcal{P}f$) we have:

$$\begin{aligned} |\mathbb{E}_{\mu_0^{\gamma_n}}(\mathcal{P}f(X_t^{\gamma_n}) - \mathcal{P}f(X_{t_k}^{\gamma_n}))| &\leq \left| \int_t^{t_k} \mathbb{E}_{\mu_0^{\gamma_n}}(\mathcal{L}_0\mathcal{P}f(X_s^\gamma)) \right| \\ &\leq (t_k - t) \sup_E |\mathcal{L}_0\mathcal{P}f| \\ &\leq \varepsilon. \end{aligned}$$

Since the bound is true for any $\varepsilon > 0$, we take $\varepsilon \rightarrow 0$ to get that Y_t equals X_t^∞ in distribution since \mathcal{D}_{eff} is assumed separating. To summarize, let $t > 0$ be given and consider any extraction of the family $(X^\gamma)_{\gamma}$. We have just proven that for each $t > 0$ we can construct a second extraction (in two steps) for which we have the convergence of the distribution of X_t^γ towards the one of X_t^∞ . This is a characteristic property of sequential convergence like convergence in distribution. \square

We will now invoke a technical Lemma that we will combine with the Markov property to prove the convergence of the finite dimensional distributions.

Lemma 3.7. *We assume as before that X^γ starts from the initial condition μ_0^γ that converges in law towards a distribution denoted μ_0 . We furthermore assume here that the support of μ_0 is included in E_{eff} , so that $\mu_0\mathcal{P} = \mu_0$. Then, when γ goes to $+\infty$, the couple (X_0^γ, X_t^γ) converges in distribution towards the couple (X_0^∞, X_t^∞) .*

Proof. We want to prove that for all $t > 0$ and f_0, f continuous bounded strictly positive functions on E , it holds:

$$\mathbb{E}_{\mu_0^\gamma}(f_0(X_0^\gamma)f(X_t^\gamma)) \xrightarrow[\gamma \rightarrow +\infty]{} \mathbb{E}_{\mu_0}(f_0(X_0^\infty)f(X_t^\infty))$$

We have that:

$$\begin{aligned} \mathbb{E}_{\mu_0^\gamma}(f_0(X_0^\gamma)f(X_t^\gamma)) &= \int_E f_0(z) \mathbb{E}_z(f(X_t^\gamma)) \mu_0^\gamma(dz) \\ &= \mu_0^\gamma(f_0) \int_E \mathbb{E}_z(f(X_t^\gamma)) \nu_\gamma(dz), \end{aligned}$$

where $\nu_\gamma(dz) = \frac{1}{\mu_0^\gamma(f_0)} f_0(z) \mu_0^\gamma(dz)$ is a probability measure on E . Since $\mu_0^\gamma(f_0)$ is uniformly lower bounded by a constant strictly superior to 0 and converges to $\mu_0(f_0)$, we have that ν_γ converges in distribution towards $\nu(dz) = \frac{1}{\mu_0(f_0)} f_0(z) \mu_0(dz)$, a probability distribution on E_{eff} (so $\nu\mathcal{P} = \nu$). Using Proposition 3.6:

$$\mathbb{E}_{\mu_0^\gamma}(f_0(X_0^\gamma) f(X_t^\gamma)) = \mu_0^\gamma(f_0) \mathbb{E}_{\nu_\gamma}(f(X_t^\gamma)) \xrightarrow{\gamma \rightarrow +\infty} \mu_0(f_0) \mathbb{E}_\nu(f(X_t^\infty)) = \mathbb{E}_{\mu_0}(f_0(X_0^\infty) f(X_t^\infty)).$$

□

We now have all the tools needed to conclude.

Proof. (Corollary 2.11)

We consider $0 < t_1 < \dots < t_k$ and $\varepsilon \in]0, t_1[$, denoting P_t^γ the semi-group associated with X_t^γ and P_t^∞ the semi-group associated with X_t^∞ , we notice that for all f_1, \dots, f_k bounded strictly positives functions on E we have:

$$\mathbb{E}_{\mu_0^\gamma}(f_1(X_{t_1}^\gamma) \dots f_k(X_{t_k}^\gamma)) = \rho_\gamma P_{h_1}^\gamma(f_1 P_{h_2}^\gamma(f_2 P_{h_3}^\gamma(\dots f_{k-1} P_{h_k}^\gamma(f_k) \dots))),$$

where $h_i = t_i - t_{i-1}$ for $i \geq 2$ and $h_1 = t_1 - \varepsilon$, and ρ_γ is the law of X_ε^γ starting from μ_0^γ . Proposition 3.6 gives us that ρ_γ converges in distribution towards ρ , the law of X_ε^∞ starting from $\mu_0\mathcal{P}$ whose support is included in E_{eff} .

To conclude, we will prove by induction on k that for any continuous bounded f

$$\nu_\gamma(f) = \rho_\gamma P_{h_1}^\gamma(f_1 P_{h_2}^\gamma(f_2 P_{h_3}^\gamma(\dots f_{k-2} P_{h_{k-1}}^\gamma(f))))$$

converges when $\gamma \rightarrow +\infty$ to:

$$\nu_\infty(f) = \rho P_{h_1}^\infty(f_1 P_{h_2}^\infty(f_2 P_{h_3}^\infty(\dots f_{k-2} P_{h_{k-1}}^\infty(f))))$$

both being positive finite non trivial measures since all the functions f_{h_i} involved are positive and bounded. Furthermore, the second measure is of support included in E_{eff} .

Proposition 3.6 proves the convergence of the semi-group which is the result for $k = 1$. We now take $k \geq 2$, we assume that the result is proved at rank $k - 1$, and use the notation ν_γ and ν_∞ for distributions at rank $k - 1$. By induction hypothesis, ν_γ converges to ν_∞ in distribution (so especially $\nu_\gamma(E)$ converges to $\nu_\infty(E)$). Writing $\overline{\nu}_\gamma := \frac{1}{\nu_\gamma(E)} \nu_\gamma$ and $\overline{\nu}_\infty := \frac{1}{\nu_\infty(E)} \nu_\infty$ we have that $\overline{\nu}_\gamma$ is a probability measure converging to $\overline{\nu}_\infty$. Hence, Lemma 3.7 gives us that for any continuous bounded function f we have:

$$\begin{aligned} \rho_\gamma P_{h_1}^\gamma(f_1 P_{h_2}^\gamma(f_2 P_{h_3}^\gamma(\dots f_{k-1} P_{h_k}^\gamma(f) \dots))) &= \nu_\gamma(E) \mathbb{E}_{\overline{\nu}_\gamma}(f_{k-1}(X_0^\gamma) f(X_{h_k}^\gamma)) \\ \xrightarrow{\gamma \rightarrow +\infty} \nu(E) \mathbb{E}_{\overline{\nu}}(f_{k-1}(X_0^\infty) f(X_{h_k}^\infty)) &= \rho P_{h_1}^\infty(f_1 P_{h_2}^\infty(f_2 P_{h_3}^\infty(\dots f_{k-1} P_{h_k}^\infty(f) \dots))). \end{aligned}$$

Since this is true for all continuous bounded functions f , this proves the result at rank k ; we take $f = f_k$ and we get that:

$$\mathbb{E}_{\mu_0^\gamma}(f_1(X_{t_1}^\gamma) \dots f_k(X_{t_k}^\gamma)) \xrightarrow{\gamma \rightarrow +\infty} \mathbb{E}_{\mu_0\mathcal{P}}(f_1(X_{t_1}^\infty) \dots f_k(X_{t_k}^\infty)).$$

□

3.6 Proof of Proposition 2.4

We propose in this section a proof of Proposition 2.4 (rewritten as Lemma 3.8 below), which states that the large time convergence of the dominant process towards E_{eff} is uniform with respect to starting points taken in compact sets.

Let us reformulate Proposition 2.4 by making assumptions explicit:

Lemma 3.8. *Let E be Polish, E_{eff} be a closed subset of E , and assume that $(x, t, \omega) \mapsto X_t(x)(\omega) \in E$ is a measurable map defined for $(t, x) \in \mathbb{R}_+ \times E$ and for ω in a probability space. A reference filtration of the considered probability space being given, assume that for all $x \in E$, $t \mapsto X_t(x)$ is adapted, càdlàg, Markovian and time homogeneous, with initial condition $X_0(x) = x$. We assume moreover that:*

1. *For all $t \geq 0$, $x \mapsto X_t(x)$ is continuous in probability (Feller-type continuity).*
2. *For all $x \in E$, the process $(X_t(x))_{t \geq 0}$ converges in probability towards $X_\infty(x) \in E_{\text{eff}}$.*
3. *$x \mapsto \mathcal{P}(x, \cdot) = \text{Law}(X_\infty(x))$ is continuous for the topology of convergence in law.*
4. *We have $\mathcal{P}(x, \cdot) = \delta_z$ for some $z \in E_{\text{eff}}$ if and only if $x \in E_{\text{eff}}$; in which case $z = x$.*
5. *The compact containment condition (2.2) is verified: for any $K \subset E$ compact and any $\varepsilon > 0$, there exists another compact $K_\varepsilon \subset E$ such that:*

$$\sup_{x \in K} \mathbb{P} [\exists t \geq 0, X_t(x) \in K_\varepsilon^c] \leq \varepsilon.$$

Then there exists a continuous bounded non-negative function f on E with $f > 0$ on $E \setminus E_{\text{eff}}$ and $f = 0$ on E_{eff} such that for all $K \subset E$ compact:

$$\lim_{t \rightarrow +\infty} \sup_{x \in K} \mathbb{E} [f(X_t(x))] = 0. \tag{3.5}$$

Let us first recall the existence of measurable projections, which will be useful in the proof of Lemma 3.8. We have the following classical result:

Lemma 3.9 (Existence of a measurable projector). *Let (E, d) denotes a Polish space, K a compact subset of E . K admits a measurable projector: there exists a measurable function $p : E \rightarrow K$ such that for all $z \in E$, $d(z, K) = d(z, p(z))$.*

Proof. We consider the set-valued map ψ on E defined by ($B(x, r)$ denotes a closed ball of center x and radius r):

$$\psi(y) := y \in E \mapsto B(y, d(y, K)) \cap K \in \mathcal{Co}(E) \setminus \{\emptyset\},$$

where $\mathcal{Co}(E)$ denotes the compact sets of E . We underscore that by construction $\psi(y)$ is always non-empty since K is compact.

For any open set $O \subset E$ we claim that

$$\mathcal{A}[O] := \{y \in E, \psi(y) \cap O \neq \emptyset\}$$

is measurable. If this claim is true, the Kuratowski-Ryll-Nardzewski measurable selection theorem (see Theorem 6.9.3 in [5]) gives us a measurable selection of the set-valued map $y \mapsto \psi(y) \subset E$, which is by construction a projector, and the proof is complete. To prove the claim, we decompose O as a countable union of closed sets

$$O = \bigcup_{n \geq 1} C_n, \quad C_n := \{x \in O \mid d(x, O^c) \geq 1/n\}$$

and write $\mathcal{A}[O] = \bigcup_{n \geq 1} \mathcal{A}[C_n]$. One can then check that since $K \cap C_n$ is compact:

$$B(y, d(y, K)) \cap K \cap C_n \neq \emptyset \Leftrightarrow d(y, K \cap C_n) \leq d(y, K).$$

The above condition is continuous with respect to the y variable, so that $\mathcal{A}[C_n]$ is closed. The claim is proved. □

To prove Lemma 3.8 we will also need to use the following general lemma. It proves, using first the continuity with respect to x of $X_t(x)$ (in probability), and second the continuity of $X_\infty(x)$ (in law), the uniform continuity of $x \mapsto X_\infty(x)$ in probability, up to some compact containment.

Lemma 3.10. *Let all the assumptions (5. excepted) stated in Lemma 3.8 hold true. Then for all $K^+ \subset E$ and $K \subset K^+$ compact sets, all $x \in K$, and all $\varepsilon, h > 0$, there exists $\eta := \eta_{K, K^+, x, \varepsilon, h} > 0$ such that for all $x' \in B(x, \eta) \cap K$, we have:*

$$\mathbb{P}(d(X_\infty(x), X_\infty(x')) > h) \leq \frac{\varepsilon}{16} + \mathbb{P}(\exists t \geq 0, X_t(x') \notin K^+) + 3\mathbb{P}(\exists t \geq 0, X_t(x) \notin K^+).$$

Proof. Before starting the proof we remark that the result is straightforward if $K^+ \cap E_{\text{eff}} = \emptyset$. We will hence assume from now on that

$$K^+ \cap E_{\text{eff}} \neq \emptyset.$$

The proof of Lemma 3.10 is then decomposed in several steps.

Two uniform continuity estimates. First, by assumption, the function $x \mapsto \mathcal{P}(x, \cdot) = \text{Law}(X_\infty(x))$ is continuous for the metric topology of convergence in law. Since K^+ is compact, Heine's theorem applies and yields uniform continuity. Remark also that for $z \in E_{\text{eff}}$ one has $\mathcal{P}(z, \cdot) = \delta_z$. As a direct consequence, for all $\alpha, h > 0$, there exists $a_{\alpha, h} > 0$ such that for all $z \in E_{\text{eff}} \cap K^+$ and $x \in K^+$:

$$d(x, z) \leq a_{\alpha, h} \quad \text{implies} \quad \mathbb{P}\left(d(X_\infty(x), z) \leq \frac{h}{2}\right) \geq 1 - \alpha. \tag{3.6}$$

Second, let $t \geq 0$ be any given time, and let us consider the continuity with respect to the initial conditions. By assumption, the map $x \mapsto X_t(x)$ is continuous for the metric topology of convergence in probability of random variables. Since K is compact, Heine's theorem applies again and yields uniform continuity, which can be expressed as follows: for all $a, \alpha > 0$, there exists $\eta_{t, a, \alpha} > 0$ such that for all $x, x' \in K$ we have

$$d(x, x') \leq \eta_{t, a, \alpha} \quad \text{implies} \quad \mathbb{P}\left(d(X_t(x), X_t(x')) \leq \frac{a}{2}\right) \geq 1 - \alpha \tag{3.7}$$

A first estimate. Let $\alpha, a > 0$ and $x \in K$ be given. Let p_{eff} denote a projector on $E_{\text{eff}} \cap K^+$ satisfying for all $x \in E$

$$p_{\text{eff}}(x) \in E_{\text{eff}} \cap K^+, \quad d(x, p_{\text{eff}}(x)) = d(x, E_{\text{eff}} \cap K^+),$$

whose existence is given by Lemma 3.9, we claim that there exists $T := T_{x, a, \alpha}$ and $\eta := \eta_{x, a, \alpha} > 0$ such that for any $x' \in B(x, \eta) \cap K$, we have:

$$\mathbb{P}(d(X_T(x'), p_{\text{eff}}(X_T(x))) \leq a) \geq 1 - 2\alpha - \mathbb{P}(\exists t \geq 0, X_t(x) \notin K^+). \tag{3.8}$$

Indeed, one has first

$$\mathbb{P}\left(d(X_T(x), E_{\text{eff}} \cap K^+) \leq \frac{a}{2}\right) \geq \mathbb{P}\left(d(X_T(x), X_\infty(x)) \leq \frac{a}{2}, X_\infty(x) \in K^+\right),$$

and since by assumption $(X_t(x))_{t \geq 0}$ converges in probability for large time to $X_\infty(x)$, using that K^+ is closed, there exists $T := T_{x, a, \alpha} > 0$ such that

$$\mathbb{P}(d(X_T(x), E_{\text{eff}} \cap K^+) \leq \frac{a}{2}) \geq 1 - \alpha - \mathbb{P}(\exists t \geq 0, X_t(x) \notin K^+). \tag{3.9}$$

Next, we can consider the lower bound

$$\begin{aligned} & \mathbb{P}(d(X_T(x'), p_{\text{eff}}(X_T(x))) \leq a) \\ & \geq \mathbb{P}\left(d(X_T(x'), X_T(x)) \leq \frac{a}{2}, d(X_T(x), p_{\text{eff}}(X_T(x))) \leq \frac{a}{2}\right) \\ & \geq 1 - \mathbb{P}\left(d(X_T(x'), X_T(x)) > \frac{a}{2}\right) - \mathbb{P}\left(d(X_T(x), E_{\text{eff}} \cap K^+) > \frac{a}{2}\right), \end{aligned}$$

and using first (3.9) for a well-chosen $T := T_{x,a,\alpha}$ and then (3.7) for a well chosen $\eta_{T,a,\alpha} := \eta_{x,a,\alpha}$ we obtain the claim (3.8).

Main estimate. The key of the proof consists in the following claim. Let $K \subset K^+ \subset E$ compact subsets, $\alpha \in [0, 1]$, $h > 0$, and $x \in K$ be given. Denote the event

$$\mathcal{B}_{x'} := \{X_t(x') \in K^+, \forall t \geq 0\}.$$

We claim that there exists $\eta > 0$ and $T \geq 0$ such that for any $x' \in B(x, \eta) \subset K$, we have:

$$\mathbb{P}\left(d(X_\infty(x'), p_{\text{eff}}(X_T(x))) \leq \frac{h}{2}\right) \geq 1 - 5\alpha - \mathbb{P}(\mathcal{B}_{x'}^c) - \mathbb{P}(\mathcal{B}_x^c).$$

This estimate will subsequently quite easily yield the proof of the whole lemma.

In order to prove the claim, we consider the following conditioning:

$$\begin{aligned} & \mathbb{P}\left(d(X_\infty(x'), p_{\text{eff}}(X_T(x))) \leq \frac{h}{2}\right) \\ & = \underbrace{\mathbb{P}(d(X_\infty(x'), p_{\text{eff}}(X_T(x))) \leq \frac{h}{2} \mid X_T(x') \in B(p_{\text{eff}}(X_T(x)), a) \cap K^+)}_{\text{ii)}} \\ & \quad \times \underbrace{\mathbb{P}(X_T(x') \in B(p_{\text{eff}}(X_T(x)), a) \cap K^+)}_{\text{i)}}. \end{aligned}$$

Using our first estimate (3.8), there exists $T := T_{x,a,\alpha} > 0$ and $\eta := \eta_{x,a,\alpha} > 0$ such that for all $x' \in B(x, \eta) \cap K$ the term i) can be lower bounded as follows:

$$\begin{aligned} \text{i)} & = \mathbb{P}(X_T(x') \in B(p_{\text{eff}}(X_T(x)), a) \cap K^+) \\ & \geq 1 - \mathbb{P}(X_T(x') \notin B(p_{\text{eff}}(X_T(x)), a)) - \mathbb{P}(\mathcal{B}_{x'}^c) \\ & \geq 1 - 2\alpha - \mathbb{P}(\mathcal{B}_x^c) - \mathbb{P}(\mathcal{B}_{x'}^c). \end{aligned}$$

Then, since $t \mapsto (X_t(x), X_t(x'))$ is homogeneous in time and satisfies the Markov property, the uniform estimate (3.6) (applied after conditioning with respect to the point $z = p_{\text{eff}}(X_t(x)) \in E_{\text{eff}} \cap K^+$ and the initial condition $X_t(x') \in K^+$) directly implies for a well chosen $a := a_{\alpha,h}$ that ii) can be lower bounded by

$$\begin{aligned} \text{ii)} & = \mathbb{P}\left(d(X_\infty(x'), p_{\text{eff}}(X_t(x))) \leq \frac{h}{2} \mid X_t(x') \in B(p_{\text{eff}}(X_t(x)), a) \cap K^+\right) \\ & \geq 1 - \alpha \end{aligned} \tag{3.10}$$

for any deterministic $t \geq 0$ and in particular for $t = T_{x,a,\alpha}$.

Finally we obtain that $\text{i)} \times \text{ii)} \geq (1 - \alpha) \times (1 - 2\alpha - \mathbb{P}(\mathcal{B}_{x'}^c) - \mathbb{P}(\mathcal{B}_x^c)) \geq 1 - 5\alpha - \mathbb{P}(\mathcal{B}_{x'}^c) - \mathbb{P}(\mathcal{B}_x^c)$ and the claim is proven.

Final remark. Eventually, $K, x \in K, K^+$, and $h, \alpha > 0$ being given as before, one can choose an appropriate T and then apply two times the main estimate in order to get for all $x' \in B(x, \eta)$:

$$\begin{aligned} & \mathbb{P}(d(X_\infty(x'), X_\infty(x)) \leq h) \\ & \geq \mathbb{P}\left(d(X_\infty(x'), p_{\text{eff}}(X_T(x))) \leq \frac{h}{2}, d(X_\infty(x), p_{\text{eff}}(X_T(x))) \leq \frac{h}{2}\right) \\ & \geq 1 - \mathbb{P}\left(d(X_\infty(x'), p_{\text{eff}}(X_T(x))) > \frac{h}{2}\right) - \mathbb{P}\left(d(X_\infty(x), p_{\text{eff}}(X_T(x))) > \frac{h}{2}\right) \\ & \geq 1 - 10\alpha - 3\mathbb{P}(\mathcal{B}_x^c) - \mathbb{P}(\mathcal{B}_{x'}^c). \end{aligned}$$

We then choose α small enough so that $10\alpha \leq \frac{\varepsilon}{16}$, and we have proved the lemma. \square

Before carrying out the proof of Lemma 3.8, we end with a classical technical lemma that will be useful in the proof. We recall as a preamble that the distance function to any closed set C is 1-Lipschitz since:

$$\inf_{y \in C} d(x', y) - \inf_{y \in C} d(x, y) \leq \sup_{y \in C} (d(x', y) - d(x, y)) \leq d(x, x').$$

Lemma 3.11. *Let (E, d) be a separable metric space. There exists a countable family of 1-Lipschitz, bounded by 1 functions that is convergence determining.*

Proof. Since E is metric separable, it is second countable and there exists a countable generating family \mathcal{B} of closed sets; meaning that for any closed set C , \mathcal{B} contains a decreasing sequence whose intersection is C . Let us consider the following countable family of bounded and Lipschitz functions

$$f_{n,C}(x) = 1 - \min(nd(x, C), 1), \quad C \in \mathcal{B}, n \in \mathbb{N} \setminus \{0\}.$$

Let C be a given closed set and $C_p \searrow C$ a decreasing sequence converging to $C \in \mathcal{B}$. One has then $\inf_{p,n} f_{n,C_p}(x) = \mathbb{1}_C(x)$. Let $(\mu_q)_q$ denotes a sequence of probabilities, and μ another probability such that $\lim_q \mu_q(f_{n,C_p}) = \mu(f_{n,C_p})$ for any n, p . Obviously, by monotone convergence

$$\limsup_q \mu_q(C) = \limsup_q \inf_{p,n} \mu_q(f_{n,C_p}) \leq \inf_{p,n} \limsup_q \mu_q(f_{n,C_p}) = \mu(C).$$

This implies by portmanteau lemma that $\lim_q \mu_q = \mu$ in distribution and the family $(f_{n,C})_{n \geq 1, C \in \mathcal{B}}$ is convergence determining.

Taking $g_{n,C_p} = f_{n,C_p}/n$ that is 1-Lipschitz and bounded by 1, we have the equivalence $\lim_q \mu_q(f_{n,C_p}) = \mu(f_{n,C_p})$ if and only if $\lim_q \mu_q(g_{n,C_p}) = \mu(g_{n,C_p})$, so the family $(g_{n,C_p})_{n \geq 1, C \in \mathcal{B}}$ is also convergence determining. \square

Proof of Lemma 3.8. We consider a given compact K of E and a given $\varepsilon > 0$. By the compact containment assumption (Assumption 3), there exists a compact K_ε such that denoting

$$\mathcal{B}_x = \{X_t(x) \in K_\varepsilon, \forall t \geq 0\}$$

we have $\sup_{x \in K} \mathbb{P}(\mathcal{B}_x^c) \leq \frac{\varepsilon}{16}$.

We take $(f_n)_{n \geq 1}$ a family of convergence determining functions on E that are bounded by 1 and 1-Lipchitz for the distance $\min(1, d)$ (Lemma 3.11). We consider for $x, y \in E$ the following (harmonic) semi-distance:

$$d_{\text{har}}(x, y) = \sum_{n \geq 1} |\mathcal{P}f_n(x) - \mathcal{P}f_n(y)| 2^{-n}.$$

Since $(f_n)_n$ is separating, we note that $d_{\text{har}}(x, y) = 0$ if and only if $\mathcal{P}(x, \cdot) = \mathcal{P}(y, \cdot)$. Moreover the continuity assumption on \mathcal{P} implies that d_{har} is topologically weaker than d .

To prove Lemma 3.8, we consider the harmonic distance function to E_{eff} defined by

$$f : x \in E \mapsto \inf_{y \in E_{\text{eff}}} d_{\text{har}}(x, y) \in \mathbb{R}_+.$$

We claim that f is null on E_{eff} , strictly positive on $E \setminus E_{\text{eff}}$ and 1-Lipschitz for d_{har} (hence continuous for d). We will subsequently prove (3.5) for this function, completing the proof of Lemma 3.8. First, we have trivially that $f|_{E_{\text{eff}}} = 0$, and that f is 1-Lipschitz as a distance function. Let us now prove that we have $f|_{E \setminus E_{\text{eff}}} > 0$. Let $x \in E \setminus E_{\text{eff}}$ be given. We assume that $f(x) = 0$ and we prove that $x \in E_{\text{eff}}$. If $f(x) = 0$, we can find by definition of f a sequence $y_n \in E_{\text{eff}}$ such that $d_{\text{har}}(y_n, x) \xrightarrow{n \rightarrow +\infty} 0$, so $\langle \delta_{y_n}, f_p \rangle = f_p(y_n) = \mathcal{P}f_p(y_n) \xrightarrow{n \rightarrow +\infty} \mathcal{P}f_p(x) = \langle \mathcal{P}(x, \cdot), f_p \rangle$ for all $p \geq 1$. Since the f_p are convergence determining, we thus have that δ_{y_n} converges weakly towards $\mathcal{P}(x, \cdot)$, but a converging sequence of Dirac distributions of support included in a closed set can only converge towards another Dirac measure of support included in the same closed set, so eventually there exists $y \in E_{\text{eff}}$ such that $\mathcal{P}(x, \cdot) = \delta_y = \mathcal{P}(y, \cdot)$. By assumption, $\mathcal{P}(x, \cdot) = \delta_y$ implies $x = y \in E_{\text{eff}}$. We have thus proved that $f|_{E \setminus E_{\text{eff}}} > 0$. We stress that this latest argument is formally important and is the main use of Assumption 5 in this paper.

We can now develop the main argument, using the first key fact that for all $n \geq 1$ and $x \in E$, the process $t \mapsto \mathcal{P}f_n(X_t(x))$ is a martingale for the underlying filtration, so that the absolute value:

$$t \mapsto |\mathcal{P}f_n(X_t(x)) - \mathcal{P}f_n(X_t(x'))|$$

is a sub-martingale whose expectation is thus increasing with time. Note also that $\mathcal{P}f_n(X_\infty(x')) = f_n(X_\infty(x'))$ almost surely since $X_\infty(x') \in E_{\text{eff}}$. The second key fact is the conclusion of Lemma 3.10, which is to be used for $K = K$, $K^+ = K_\varepsilon$, $x = x$, $h = \frac{\varepsilon}{8}$, a well-chosen $\eta_{x,\varepsilon} := \eta_{K, K_\varepsilon, x, \varepsilon}$ and all $x' \in B(x, \eta_{x,\varepsilon}) \cap K$.

We then proceed to estimate $\mathbb{E}(f(X_t(x')))$ as follows:

$$\begin{aligned} \mathbb{E}(f(X_t(x'))) &= |\mathbb{E}(f(X_t(x')) - f(X_\infty(x')))| \\ &\leq \mathbb{E}(d_{\text{har}}(X_t(x'), X_\infty(x'))) \\ &\leq \sum_{n \geq 1} 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_t(x')) - \mathcal{P}f_n(X_t(x))|) \\ &\quad + 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_t(x)) - \mathcal{P}f_n(X_\infty(x))|) \\ &\quad + 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_\infty(x)) - \mathcal{P}f_n(X_\infty(x'))|). \end{aligned}$$

Applying the sub-martingale property on the first term of the r.h.s., it yields

$$\begin{aligned} \mathbb{E}(f(X_t(x'))) &\leq \sum_{n \geq 1} 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_t(x)) - \mathcal{P}f_n(X_\infty(x))|) \\ &\quad + 2^{-n} \mathbb{E}(\underbrace{|\mathcal{P}f_n(X_\infty(x)) - \mathcal{P}f_n(X_\infty(x'))|}_{|f_n(X_\infty(x)) - f_n(X_\infty(x'))|}) \times 2 \\ &\leq \sum_{n \geq 1} 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_t(x)) - \mathcal{P}f_n(X_\infty(x))|) \\ &\quad + 2^{-n} \mathbb{E}(d(X_\infty(x), X_\infty(x')) \wedge 1) \times 2 \end{aligned}$$

where in the last line we have used the Lipschitz bound on the functions f_n . Next, using

first Lemma 3.10 and then the property of \mathcal{B}_x defined in the beginning of this proof :

$$\begin{aligned} \mathbb{E}(f(X_t(x'))) &\leq \sum_{n \geq 1} 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_t(x)) - \mathcal{P}f_n(X_\infty(x))|) \\ &\quad + 2^{-n} \left(\frac{\varepsilon}{8} + \mathbb{P}(d(X_\infty(x), X_\infty(x')) > \frac{\varepsilon}{8}) \right) \times 2 \\ &\leq \sum_{n \geq 1} 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_t(x)) - \mathcal{P}f_n(X_\infty(x))|) \\ &\quad + 2^{-n} \left(\frac{3\varepsilon}{16} + 3\mathbb{P}(\mathcal{B}_x^c) + \mathbb{P}(\mathcal{B}_{x'}^c) \right) \times 2 \\ &\leq \sum_{n \geq 1} 2^{-n} \mathbb{E}(|\mathcal{P}f_n(X_t(x)) - \mathcal{P}f_n(X_\infty(x))|) \\ &\quad + \frac{7\varepsilon}{8} \end{aligned}$$

Using that the f_n are bounded by 1, the first term above converges to 0 when t approaches infinity, and is thus inferior to $\frac{\varepsilon}{8}$ for all $t > T_{x,\varepsilon}$ for some $T_{x,\varepsilon}$.

By covering K by a finite number of balls (we can because K is compact) of centers x_i radius $\eta_{x_i,\varepsilon}$ for $i = 1 \dots I_\varepsilon$, we thus find that for all $t \geq \max_i T_{x_i,\varepsilon}$ it holds

$$\sup_{x' \in K} \mathbb{E}[f(X_t(x'))] \leq \varepsilon.$$

Since this is true for all $\varepsilon > 0$, we have eventually:

$$\lim_{t \rightarrow +\infty} \sup_{x \in K} \mathbb{E}[f(X_t(x))] = 0.$$

□

4 Examples

We will now use Theorem 2.10 and Corollary 2.11 on concrete examples to get homogenization results. Once the framework is settled, we only have to check the common hypothesis of both Theorem 2.10 and Corollary 2.11 to have the convergence in law respectively in the sense of the Meyer–Zheng topology and in the sense of the finite-dimensional distributions. Since we will study SDEs with bounded, Lipschitz coefficients, Lemma B.2 already gives us Assumptions 4 and Assumption 6.

4.1 Diffusion in a simplex

This example is directly taken from [13].

Consider E a closed simplex in \mathbb{R}^n , that is the non-degenerate intersection of $n + 1$ affine half-spaces E_i , $i = 1 \dots n + 1$. The effective state space E_{eff} is defined to be the $n + 1$ vertices of E : we have therefore that E is compact and E_{eff} is a finite set. We take $\mathcal{D}_{\text{eff}} = \mathcal{C}_c(E_{\text{eff}})$ which is here the set of all real functions on E_{eff} . We assume in this example that the dominant process is a martingale and that the subdominant process is general, we write:

$$dX_t^\gamma = \sqrt{\gamma} \sigma(X_t^\gamma) dW_t + \sigma_0(X_t^\gamma) dB_t + b(X_t^\gamma) dt,$$

where we assume that $\sigma, \sigma_0 : E \rightarrow \mathcal{M}_n(\mathbb{R})$ and $b : E \rightarrow \mathbb{R}^n$ are Lipschitz continuous (for the Euclidean distance). Thus we have:

$$\mathcal{L}_0 = \langle b, \nabla \rangle + \frac{1}{2} \langle \sigma_0 \sigma_0^\dagger \nabla, \nabla \rangle \quad \text{and} \quad \mathcal{L}_1 = \frac{1}{2} \langle \sigma \sigma^\dagger \nabla, \nabla \rangle$$

In order to obtain an appropriate behaviour of the above process at the boundary ∂E , it is necessary to add several assumptions:

- (i) For all $x \in \partial E$, $\sigma(x)$ and $\sigma_0(x)$ have their images in the tangent space of ∂E_i for each i with $x \in \partial E_i$. In particular $\sigma(x) = \sigma_0(x) = 0$ if $x \in E_{\text{eff}}$.
- (ii) For all $x \in \partial E$, $b(x)$ vanishes or points to the interior of E .
- (iii) $\sigma(x) = 0$ if and only if $x \in E_{\text{eff}}$.

The above special assumptions ensure in particular that the process stays in E almost surely for any $\gamma > 0$ and $t \geq 0$ (see [13]).

Lemma 4.1. *Assume that in addition to Lipschitz continuity, σ , σ_0 and b satisfy i) and ii) above. Then X_t^γ belongs to E for all time t .*

We will next prove again the main result of [13], that is:

Theorem 4.2. *The process $X^{(\gamma,1)}$ converges in law (for the pseudo-paths topology and for finite dimensional time marginals) when γ approaches infinity towards a continuous-time Markov chain X^∞ on E_{eff} of transition matrix $\mathcal{L}_{\text{eff}} = (b(x) \cdot \nabla_x H_z(x))_{x,z \in E_{\text{eff}}}$, where $x \mapsto H_z(x)$ is the only affine function on E such that: i) $H_z(z) = 1$, and ii) $H_z(z') = 0$ for any $z' \in E_{\text{eff}} \setminus \{z\}$.*

We just have to check the six remaining hypothesis.

Since the dominant process is a bounded martingale, its quadratic variation is almost surely finite so that the dominant process converges almost surely to a point where σ is null; by iii) above the latter is a point of E_{eff} . $X_t^{(1,0)}$ converges in probability towards E_{eff} when $t \rightarrow +\infty$, and we have $X_\infty^{(1,0)}(x) = x$ if $x \in E_{\text{eff}}$: Assumption 1 is checked.

The assumption 3 is trivially verified since E is a compact.

Let us now give an explicit expression of \mathcal{P} to prove the remaining hypothesis. It has been proved in [13], Remark 6, that for all $x \in E$:

$$\mathcal{P}(x, \cdot) = \sum_{z \in E_{\text{eff}}} H_z(x) \delta_z,$$

In short, the proof works as follows: $X_\infty^{(1,0)}(x) \in E_{\text{eff}}$ and $X^{(1,0)}$ is a martingale, hence taking the expected value of $X_\infty^{(1,0)}(x)$ we have that x is a convex combination of the $n+1$ points of E_{eff} weighted by the $\mathbb{P}(X_\infty(x) = z)$, and it implies $\mathbb{P}(X_\infty(x) = z) = H_z(x)$.

The explicit expression of \mathcal{P} immediately gives us Assumption 5.

The Assumption 2 and Assumption 7 are easily proved using the fact that \mathcal{P} is affine: $\mathcal{P}\varphi$ being smooth on E for all $\varphi \in \mathcal{D}_{\text{eff}}$, point ii) is a direct consequence of Itô's formula applied on $\mathcal{P}\varphi(X^{(\gamma,1)})$. Indeed, for all $\varphi \in \mathcal{D}_{\text{eff}}$, we have:

$$\begin{aligned} d(\mathcal{P}\varphi(X_t^{(\gamma,1)})) &= (\gamma \mathcal{L}_1 + \mathcal{L}_0)(\mathcal{P}\varphi)(X_t^{(\gamma,1)})dt + \nabla(\mathcal{P}\varphi)(X_t^\gamma) \cdot \sigma_0(X_t^{(\gamma,1)})dB_t \\ &\quad + \sqrt{\gamma} \nabla(\mathcal{P}\varphi)(X_t^\gamma) \cdot \sigma(X_t^{(\gamma,1)})dW_t, \end{aligned}$$

integrating this inequality and using that $\mathcal{L}_1 \mathcal{P}\varphi = 0$, we get eventually:

$$\begin{aligned} \mathcal{P}\varphi(X_t^{(\gamma,1)}(x)) &- \int_0^t \mathcal{L}_0 \mathcal{P}\varphi(X_s^{(\gamma,1)}(x))ds \\ &= \mathcal{P}\varphi(x) + \int_0^t \nabla(\mathcal{P}\varphi)(X_s^{(\gamma,1)}(x)) \cdot \sigma_0(X_s^{(\gamma,1)}(x))dB_s \\ &\quad + \int_0^t \sqrt{\gamma} \nabla(\mathcal{P}\varphi)(X_s^{(\gamma,1)}(x)) \cdot \sigma(X_s^{(\gamma,1)}(x))dW_s, \end{aligned}$$

with the second term clearly being a martingale for the natural filtration of $X_t^{(\gamma,1)}(x)$.

For point *iii*), using the explicit expression of $\mathcal{L}_0\mathcal{P}$ and the continuity of b and ∇H_z , we have for $z_0 \in E_{\text{eff}}$ that:

$$\lim_{x \rightarrow z_0} \mathcal{L}_0\mathcal{P}[\varphi](x) = \lim_{x \rightarrow z_0} \sum_{z \in E_{\text{eff}}} \varphi(z)b(x) \cdot \nabla H_z(x) = \sum_{z \in E_{\text{eff}}} \varphi(z)b(z_0) \cdot \nabla H_z(z_0).$$

Identifying φ with the vector $(\varphi(z))_{z \in E_{\text{eff}}}$ we have that $\mathcal{L}_{\text{eff}}\varphi = (b(x) \cdot \nabla_x H_z(x))_{x, z \in E_{\text{eff}}}$ φ , which is trivially bounded measurable.

Since $b(x_0) \cdot \nabla H_{x_0}(x_0) \leq 0$ and $b(x_0) \cdot \nabla H(x_0) \geq 0$ for $z \neq x_0$, using furthermore that $\sum_{z \in E_{\text{eff}}} b(z) \cdot \nabla H_z(x_0) = 0$, we can check that \mathcal{L}_{eff} is indeed the generator of a continuous-time Markov process on E_{eff} . The martingale problem is therefore well-posed and Assumption 8 is satisfied.

4.2 Diffusion in a strip

Define $E = [-1, 1] \times \mathbb{R}$ and E_{eff} its border $\{-1, 1\} \times \mathbb{R}$. We take $\mathcal{D}_{\text{eff}} = C_c^\infty(\{-1, 1\} \times \mathbb{R})$ as space of test functions. In this subsection we will study the process $X^\gamma = X^{(\gamma, 1)} = (Y^\gamma, Z^\gamma)$ defined by:

$$\begin{cases} dY_t^\gamma = \sqrt{\gamma}\sigma(Y_t^\gamma, Z_t^\gamma) dW_t + b(Y_t^\gamma, Z_t^\gamma)dt \\ dZ_t^\gamma = \gamma\sigma(Y_t^\gamma, Z_t^\gamma)^2 dt = d[Y^\gamma]_t. \end{cases}$$

where $b : [-1, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ is a Lipschitz bounded function such that $b(1, z) \leq 0$ and $b(-1, z) \geq 0$ and $\sigma : [-1, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ is a Lipschitz bounded function with $\sigma^{-1}(\{0\}) = \{\pm 1\} \times \mathbb{R}$. Note that when the values of b and σ do not depend of z -coordinate, the above process is the simplex example of Section 4.1 for $n = 1$ (the variable Y) with the addition of its quadratic variation (the variable Z). Adding the quadratic variation is of special interest to keep track of the "size" of the process trajectory through the homogenization procedure.

By setting $b(y, z) := b(\frac{y}{|y|}, z)$ and $\sigma(y, z) := 0$ when $y \in \mathbb{R} \setminus [-1, 1]$, we can extend b and σ to Lipschitz bounded coefficients on \mathbb{R}^2 . Defining $E = \{(y, z) : \xi(y, z) \leq 0\}$ with the level function $\xi(y, z) = y^2 - 1$, one can check that the assumptions of Lemma B.1 are satisfied so that strong solutions of the SDE above indeed remain in E for all times.

The main homogenization theorem is then the following.

Theorem 4.3. *The process $X^{(\gamma, 1)}$ in $[-1, 1] \times \mathbb{R}_+$ converges in law (pseudo-path topology and finite-dimensional) when γ approaches infinity towards a Levy-type process of generator:*

$$\begin{aligned} \mathcal{L}_{\text{eff}}(\varphi)(y, z) = & |b|(y, z) \int_{\mathbb{R}_+} [\varphi(y, z+h) - \varphi(y, z)]d\mu(h) \\ & + |b|(y, z) \int_{\mathbb{R}_+} [\varphi(-y, z+h) - \varphi(y, z)]d\nu(h) \end{aligned}$$

in which ν is a finite positive measure of total mass $\nu(\mathbb{R}_+) = 1/2$, and μ a positive Levy measure of a subordinator ($\int_{\mathbb{R}_+} \min(1, t)\mu(dt) < +\infty$). The latter are given by:

$$\begin{cases} \mu(dt) := -\lim_{y \rightarrow 1} \frac{\partial}{\partial y} \mathbb{P}_y(\tau \in dt; W_\tau = 1) = \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 e^{-t \frac{\pi^2 k^2}{8}} dt \\ \nu(dt) := -\lim_{y \rightarrow 1} \frac{\partial}{\partial y} \mathbb{P}_y(\tau \in dt; W_\tau = -1) = -\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} dt, \end{cases}$$

in which $(W_t)_{t \geq 0}$ denotes a standard Brownian motion and τ its hitting time of ± 1 .

We have to check the six remaining hypothesis.

On the one hand we have that $(Y_t^{(1,0)})_t$ is a bounded martingale that converges almost surely to $Y_\infty^{(1,0)} \in \{\pm 1\}$. On the other hand, since $Y_t^{(1,0)} = \int_0^t \sigma(Y_s^{(1,0)}, Z_s^{(1,0)}) dW_s$, we have using Itô isometry and $|Y_t| \leq 1$ that:

$$\int_0^t \mathbb{E}_{(y,z)} \left(\sigma(Y_s^{(1,0)}, Z_s^{(1,0)})^2 \right) ds = \mathbb{E}_{(y,z)} \left(\left(\int_0^t \sigma(Y_s^{(1,0)}, Z_s^{(1,0)}) dW_s \right)^2 \right) \leq 1.$$

The left hand side above is exactly the expected value of $Z_t^{(1,0)} - z$ and since $(Z_t^{(1,0)})_t$ is increasing, we may take the limit in the last inequality to find that $\mathbb{E}(Z_\infty^{(1,0)}) \leq z + 1$. This implies in particular that $Z_\infty^{(1,0)} < +\infty$ almost surely. Finally, $(Y_t^{(1,0)}, Z_t^{(1,0)})$ converges almost surely when $t \rightarrow +\infty$ towards a random variable $(Y_\infty^{(1,0)}, Z_\infty^{(1,0)})$ and we have $(Y_\infty^{(1,0)}, Z_\infty^{(1,0)})(y, z) = (y, z)$ if $(y, z) \in E_{\text{eff}}$: Assumption 1 is verified.

Next, we recall that:

$$\begin{aligned} Y_t^\gamma &= \sqrt{\gamma} \int_0^t \sigma(Y_s^\gamma, Z_s^\gamma) dW_s + \int_0^t b(Y_s^\gamma, Z_s^\gamma) ds + y \\ &= \sqrt{\gamma} M_t + B_t, \end{aligned}$$

in which Y_t^γ is bounded in absolute value by 1 and B_t by $t \times \|b\|_\infty + 1$; hence Itô's isometry gives us for all $t \geq 0$ and $\gamma \geq 0$:

$$\sup_{(y,z)} \mathbb{E}_{(y,z)} \left(\gamma \int_0^t \sigma(Y_s^\gamma, Z_s^\gamma)^2 ds \right) \leq (1 + t \times \|b\|_\infty + 1)^2.$$

The term inside the expectation above is exactly $Z_t^\gamma - z$, so using Markov's inequality, we have that for all $\varepsilon > 0$, there exists $M_{\varepsilon,t} > 0$ uniform in $\gamma > 0$ such that:

$$\sup_{(y,z)} \mathbb{P}_{(y,z)} (|Z_t^\gamma - z| > M_{\varepsilon,t}) \leq \varepsilon.$$

Taking $K_{\varepsilon,T} = [-1, 1] \times [-z_0, z_0 + M_{\varepsilon,T}]$ for $z_0 > 0$ and using that Z_t^γ is increasing in time, we have for any z_0 :

$$\sup_{y \in [-1, 1], -z_0 \leq z \leq z_0} \mathbb{P}_{(y,z)} (\exists t \in [0, T], (Y_t^\gamma, Z_t^\gamma) \in K_{\varepsilon,T}^c) \leq \varepsilon.$$

The proof works in particular for $b = 0$ (when we only consider the dominant process). The estimate being uniform in $\gamma > 0$, the compact containment condition in Assumption 3 is thus checked.

To prove the remaining assumptions, Assumption 2, Assumption 5, Assumption 7 and Assumption 8, we first need to have an explicit expression of \mathcal{P} and its derivative with respect to y .

We start with a time change: the process of generator \mathcal{L}_1/σ^2 is of the form:

$$\begin{cases} d\tilde{Y}_t = dW_t \\ d\tilde{Z}_t = dt. \end{cases}$$

and we have $(\tilde{Y}_{S_t}, \tilde{Z}_{S_t}) = (Y_t, Z_t)$ where $\int_0^t \sigma^{-2}(Y_s, Z_s) ds = S_t$. On the one hand if we write $\tau := \tau_{-1,1}(W)$ the first hitting time of 1 or -1 by the Brownian motion W starting at y , we have that $\tilde{Y}_\tau \in \{\pm 1\}$ and $\tilde{Y}_t \in]-1, 1[$ for all $t < \tau$. Since $Y_t \in]-1, 1[$ for all $t \geq 0$

and $Y_\infty = \{\pm 1\}$, we can write $S_\infty = \tau$ and obtain $Y_\infty = \tilde{Y}_\tau = W_\tau$ and thus $Z_\infty = \tilde{Z}_\tau = \tau$. Eventually we have for any $\varphi : E_{\text{eff}} \rightarrow \mathbb{R}$:

$$\mathcal{P}(\varphi)(y, z) = \mathbb{E}_{(y,z)}(\varphi(Y_\infty, Z_\infty)) = \mathbb{E}_y[\varphi(W_\tau, z + \tau)]$$

As a consequence,

$$\begin{aligned} \mathcal{P}\varphi(y, z) &= \int_0^{+\infty} \varphi(1, z + t) \mathbb{P}_y(\tau \in dt; W_\tau = 1) \\ &\quad + \int_0^{+\infty} \varphi(-1, z + t) \mathbb{P}_y(\tau \in dt; W_\tau = -1), \end{aligned}$$

where $\mathbb{P}_y(\tau \in dt; W_\tau = \pm 1)$ is a transparent notation for the probability distribution of (τ, W_τ) . This explicit expression of \mathcal{P} and a dominated convergence argument gives us easily that if $\varphi \in C_b(E_{\text{eff}})$, $\mathcal{P}\varphi$ is continuous on E with limit φ at E_{eff} and hence Assumption 2 is verified. The explicit expression of \mathcal{P} furthermore gives us immediately Assumption 5. Standard regularity results on parabolic PDE (see [12, Ch. 7]) gives us that $(y, z) \mapsto \mathcal{P}\varphi(y, z)$ is in $\mathcal{C}^\infty(E)$ (*Nota Bene*: this also proves again Assumption 2). For the sake of completeness, we give in appendix an elementary proof of the fact that $\mathcal{P}\varphi$ is infinitely differentiable on E using an explicit expression of $\mathbb{P}_y(\tau \in dt; W_\tau = 1)$. We can thus use Itô's formula on to get Assumption 7 point *ii*).

We denote μ, ν the two measures defined for $t > 0$ by:

$$\begin{cases} \mu(dt) &= -\lim_{y \rightarrow 1} \frac{\partial}{\partial y} \mathbb{P}_y(\tau \in dt; W_\tau = 1) \\ \nu(dt) &= -\lim_{y \rightarrow 1} \frac{\partial}{\partial y} \mathbb{P}_y(\tau \in dt; W_\tau = -1) \end{cases}$$

and for $t \leq 0$ by $\mu(dt) = \nu(dt) = 0$. We prove in Appendix C that these two measures are positive, and are in fact Levy measures of subordinators. We are looking for an explicit expression of $\mathcal{L}_{\text{eff}}\varphi(1, z)$ using μ and ν , but we cannot directly take the limit under the integral when y approaches 1 in $\partial_y \mathcal{P}\varphi$ since $\int \varphi(t, z) \mu(dt)$ could be infinite. We thus need to subtract to $\varphi(y', z')$ the constant $\varphi(1, z)$ and obtain

$$\lim_{y \rightarrow 1} \mathcal{L}_0 \mathcal{P}\varphi(y, z) = \lim_{y \rightarrow 1} b(y, z) \partial_y \mathcal{P}(\varphi - \varphi(1, z))(y, z)$$

We prove in appendix using a standard dominated convergence argument that we can take the following limit under the integral, b being continuous by assumption:

$$\begin{aligned} \lim_{y \rightarrow 1} \mathcal{L}_0 \mathcal{P}\varphi(y, z) &= -b(1, z) \int_0^{+\infty} (\varphi(1, z + t) - \varphi(1, z)) \mu(dt) \\ &\quad - b(1, z) \int_0^{+\infty} (\varphi(-1, z + t) - \varphi(1, z)) \nu(dt) \end{aligned} \tag{4.1}$$

where the last two integrals are well-defined and uniformly bounded in z since we have $\varphi \in \mathcal{C}_c^\infty(\{-1, 1\} \times \mathbb{R})$ and $1 \wedge |x|$ integrable for μ and ν is finite.

Remark 4.4. We remind that $b(1, z) \leq 0$ (the process stays in E) and therefore that $-b(1, z) \geq 0$.

Doing exactly the same calculus as above, but taking $y \rightarrow -1$ instead of $y \rightarrow 1$, we find:

$$\begin{aligned} \lim_{y \rightarrow -1} \mathcal{L}_0 \mathcal{P}\varphi(y, z) &= b(-1, z) \int_0^{+\infty} (\varphi(1, z + t) - \varphi(-1, z)) \nu(dt) \\ &\quad + b(-1, z) \int_0^{+\infty} (\varphi(-1, z + t) - \varphi(-1, z)) \mu(dt) \end{aligned} \tag{4.2}$$

Combining (4.1) and (4.2) and using that $\mu(1 \wedge |x|), \nu(1) < +\infty$ we get that $(\pm 1, z) \mapsto \mathcal{L}_{\text{eff}} \varphi(\pm 1, z)$ is a bounded measurable function on E_{eff} and we have therefore that Assumption 7 *iii*) is checked. Furthermore \mathcal{L}_{eff} is a Levy generator constructed using homogeneous subordinators on \mathbb{R}_+ : the martingale problem is well-posed and Assumption 8 is checked. A slightly general reference that enables to obtain well-posedness of martingale problems associated with Levy generators is Theorem 3.1 of [17].

4.3 Diffusion in an Euclidean ball

We define $E = \overline{\mathbb{B}^n} \subset \mathbb{R}^n$ the closed unit ball of \mathbb{R}^n and $E_{\text{eff}} = \mathbb{S}^{n-1} = \partial \mathbb{B}^n$ its border, the unit sphere. We take $\mathcal{D}_{\text{eff}} = C_c^\infty(\mathbb{S}^{n-1})$ which is the set of all smooth functions on E_{eff} as space of test functions. We consider a purely diffusive dominant process on the closed unit ball $\overline{\mathbb{B}^n}$ of \mathbb{R}^n and a repulsive drift on the sphere for the sub-dominant one:

$$dX_t^\gamma = \sqrt{\gamma} \sigma(X_t^\gamma) dW_t + b(X_t^\gamma) dt,$$

where $\sigma : \overline{\mathbb{B}^n} \rightarrow \mathbb{R}_+$ is Lipschitz and vanishes *exactly* on $\partial \overline{\mathbb{B}^n} = \mathbb{S}^{n-1}$, while $b : \overline{\mathbb{B}^n} \rightarrow \mathbb{R}^n$ is Lipschitz and points towards the interior of the ball at points of \mathbb{S}^{n-1} ($x \cdot b(x) \leq 0$). Our process is thus a strong solution to a SDE with Lipschitz coefficients. Moreover, by extending σ and b on \mathbb{R}^n to Lipschitz functions with $\sigma(x) = 0$ and $x \cdot b(x) \leq 0$ for x outside the closed ball, we readily obtain:

$$\frac{1}{2} \gamma \sigma(x)^2 \Delta_x(|x|^2 - 1) + b(x) \cdot \nabla_x(|x|^2 - 1) \leq 0, \quad x \notin \overline{\mathbb{B}^n}$$

According to Lemma B.1, the process remains in E for all time.

We will denote by

$$b_r(x) := \langle b(x), \frac{x}{|x|} \rangle \in \mathbb{R},$$

and

$$b_\theta(x) := b(x) - b_r(x) \frac{x}{|x|} \in \mathbb{R}^n,$$

the polar decomposition of b . We also denote by

$$T\varphi(x) := (\nabla_x - \frac{x}{|x|} \langle \frac{x}{|x|}, \nabla_x \rangle) \varphi(x) \in T_x \mathbb{S}^{n-1},$$

the (tangent-valued) gradient of a smooth map on the unit sphere $\varphi : \mathbb{S}^{n-1} \rightarrow \mathbb{R}$ evaluated at $x \in \mathbb{S}^{n-1}$.

Theorem 4.5. *The process $X^{(\gamma,1)}$ with values in $\overline{\mathbb{B}^n}$ converges in law (with respect to the pseudo-path topology and finite-dimensional marginals) when γ approaches infinity towards a Levy-type jump process of generator:*

$$\begin{aligned} \mathcal{L}_{\text{eff}}(\varphi)(x) = & b_\theta(x) \cdot T\varphi(x) \\ & - 2b_r(x) \int_{\mathbb{S}^{n-1}} (\varphi(y) - \varphi(x) - \langle T\varphi(x), y - x \rangle) d\mu_x(y). \end{aligned}$$

where $d\mu_x(y) = \frac{1}{\|x-y\|^n} d\sigma^{n-1}(y)$ is a Levy measure on the sphere equipped with the uniform probability measure σ^{n-1} .

We just have to check the six other hypothesis.

Since the dominant process in a bounded martingale it converges almost surely to a point $X_\infty^{(1,0)}$ by Doob's first martingale convergence theorem. By computing its quadratic variation, which must be finite by Itô isometry, the limit $\sigma(X_\infty^{(1,0)})$ is null and therefore lies in E_{eff} . $X_t^{(1,0)}$ converges when $t \rightarrow +\infty$ towards $X_\infty^{(1,0)} \in E_{\text{eff}}$ and $X_\infty(x) = x$ if $x \in E_{\text{eff}}$: Assumption 1 is checked.

Assumption 3 is trivially verified since $\overline{\mathbb{B}^n}$ is a compact.

In an Euclidean ball, we do have an explicit expression (see [9, Ch.II Sec. 1]) of the harmonic measure:

$$\mathcal{P}(x, \cdot) = \begin{cases} \frac{1-\|x\|^2}{\|x-y\|^n} d\sigma^{n-1}(y) & \text{if } x \in \mathbb{B}^n \\ \delta_x & \text{if } x \in \mathbb{S}^{n-1} \end{cases},$$

where σ^{n-1} is the uniform probability measure on \mathbb{S}^{n-1} . This expression immediately gives us Assumption 5. $x \mapsto \mathcal{P}\varphi(x)$ is continuous on $\overline{\mathbb{B}^n}$ (we let the proof in appendix for points of \mathbb{S}^{n-1}) and we get Assumption 2.

Elliptic regularity immediately gives us that if φ is smooth on the unit sphere, then $x \mapsto \mathcal{P}\varphi(x)$ is also smooth on the unit ball, see [12, Ch. 6] (*Nota bene*: this also proves again Assumption 2). We may thus use Itô's formula to get Assumption 7 point *ii*).

We now consider the non-radial derivatives. Let us denote by D_θ the projection on the orthogonal space of $z \in \mathbb{R}^n$ of the gradient at the point z . The C^1 regularity of $\mathcal{P}\varphi$ on the unit ball yields:

$$\lim_{z \rightarrow a} D_\theta \mathcal{P}\varphi(z) = T\mathcal{P}\varphi(a) = T\varphi(a) \tag{4.3}$$

since $\varphi = \mathcal{P}\varphi$ on the unit sphere.

Let us now denote by ∂_r the projection on the vector line generated by $z \in \mathbb{R}^n$ of the gradient at the point z . A computation detailed in Appendix D gives us that:

$$\lim_{z \rightarrow a} \frac{\partial}{\partial r} \mathcal{P}\varphi(z) = -2 \int_{\mathbb{S}^{n-1}} \frac{1}{\|a-y\|^n} (\varphi(y) - \varphi(a) - \langle T\varphi(a), y-a \rangle) d\sigma^{n-1}(y), \tag{4.4}$$

where we notice that, writing $d\mu_a(y) := \frac{1}{\|a-y\|^n} d\sigma^{n-1}(y)$, we have:

$$\int_{\mathbb{S}^{n-1}} (1 \wedge \|a-y\|^2) d\mu_a(y) < +\infty \tag{4.5}$$

Thus, combining (4.3) and (4.4) and, we have that:

$$\begin{aligned} \lim_{z \rightarrow a} \mathcal{L}_0 \mathcal{P}\varphi(z) &= -2b_r(a) \int_{\mathbb{S}^{n-1}} (\varphi(y) - \varphi(a) - \langle T\varphi(a), y-a \rangle) d\mu_a(y) \\ &\quad + b_\theta(a) \cdot T\varphi(a). \end{aligned}$$

Combining the facts that b and all derivatives of φ are bounded and (4.5), we have that $\mathcal{L}_{\text{eff}}\varphi$ is measurable bounded on \mathbb{S}^{n-1} and we have point *iii*) of Assumption 7.

In the case where $b_\theta = 0$ on the one hand and $b_r(x)$ is independent of x on the other hand, the explicit expression of \mathcal{L}_{eff} shows that it is the generator of a homogeneous (rotation invariant) Levy process on \mathbb{S}^{n-1} : the martingale problem is therefore well-posed and Assumption 8 is checked. The well-posedness for general Lipschitz coefficients b_θ, b_r could then be carried out directly using a drift and a time change that map the martingale solution to the homogeneous case. It is nonetheless not so easy to find general references for this. The well-posedness can rather be obtained in regular cases like here with Theorem 3.1 of [17]. The latter theorem requires to write the considered Levy measure, here $\mu_x(dy)$, as the image by a mapping (at least Lipschitz with respect to x) of the form $y = \gamma(x, u)$ of a reference infinite measure integrating the variable u independently of x . This can be done in our case using rotational invariance by simply setting $\gamma(x, \rho, m) = \exp_x(\rho P_x m)$ in which \exp is the usual exponential map, $P_x m$ is the projection on $T_x \mathbb{S}^{n-1}$ of a unit vector $m \in \mathbb{R}^n$ uniformly integrated, and $\rho \in]0, \pi[$ is integrated with a Levy measure. The mappings $x \mapsto \gamma(x, \rho, m) = \exp_x(\rho P_x m)$, $x \mapsto b_r(x)$ and $x \mapsto b_\theta(x)$ are Lipschitz with uniform Lipschitz constants so that the various Lipschitz conditions in Theorem 3.1 of [17] are satisfied.

Appendices

Appendix A The pseudo-paths topology and the Meyer-Zheng criterion

All results presented in this section can be found in the seminal work of Meyer and Zheng [18], up to minor presentation variations.

A.1 The pseudo-path topology

Let E be a Polish space. Two measurable paths $t \mapsto x_t \in E$ and $t \mapsto y_t \in E$ belong to the same equivalence class, called a *pseudo-path*, if the set $\{t : x_t = y_t\}$ is of full Lebesgue measure (similarly as functions in \mathbb{L}^p). Let us denote by $L^0(e^{-t}dt, E)$ the set of all pseudo-paths. The *pseudo-path topology* is the Polish topology on $L^0(e^{-t}dt, E)$ induced by the closed (see IV.43 [8] for a proof) injection of $L^0(e^{-t}dt, E)$ in $\Pr(E \times [0, +\infty])$ endowed with the usual topology of convergence in law. This injection maps any pseudo-path x to the probability measure of the form

$$\delta_{x_t}(dx)e^{-t}dt. \quad (\text{A.1})$$

In probabilistic terms, x^n converges to x for the pseudo-path topology if and only if the pair (x_T^n, T) converges in distribution towards (x_T, T) , where T is exponentially distributed.

By a simple continuity argument, the push-forward (A.1) from paths to pseudo-paths is injective on the set $\mathbb{D}(\mathbb{R}_+, E)$ of càdlàg paths, so that there exists a natural injection $\mathbb{D}(\mathbb{R}_+, E) \subset L^0(e^{-t}dt, E)$. It is important to keep in mind that the set of pseudo-paths with a càdlàg representative is a Borel subset (it is in fact G_δ) but it is *not* closed in the space of pseudo-paths. As a consequence if a sequence of càdlàg paths converges for the pseudo-paths topology, the limit may or may not be càdlàg.

There is another natural way to define a Polish topology on the space of pseudo-paths $L^0(e^{-t}dt, E)$ by considering convergence in measure (equivalent here to convergence in probability). This topology can be defined with a probabilistic perspective as follows. A sequence of pseudo-paths x^n converges in probability (or in measure) towards x if and only if x_T^n converges in probability towards x_T were T is exponentially distributed. Convergence in probability on $L^0(e^{-t}dt, E)$ is then a Polish topology that can classically be metrized using the \mathbb{L}^1 -type Ky Fan complete metric (see [10, Section 9.2])

$$\int_0^\infty \min(d(x_t, y_t), 1)e^{-t}dt = 0 \quad (\text{A.2})$$

where d is complete and metrizing E . The following lemma is a slight generalization of Lemma 1 in [18] which is proved for real valued processes. It shows that the above two Polish topologies on pseudo-paths are topologically equivalent (although not metrically).

Lemma A.1. *Let E denote a Polish space. On pseudo-paths space $L^0(e^{-t}dt, E)$, the Polish topology of convergence in probability induced by (A.2) is equivalent to the Polish topology of convergence in distribution induced by (A.1).*

Proof. Let us prove that convergence in measure for pseudo-paths implies convergence as distributions on $\mathbb{R}_+ \times E$. Let T be an exponentially distributed random variable. By definition of convergence in measure, x_T^n converges in probability towards x_T ; and so is the pair (x_T^n, T) towards (x_T, T) . This routinely implies convergence in distribution in $\Pr(E \times [0, +\infty])$.

Conversely, let us prove that convergence $x^n \rightarrow x$ as probability distributions on $\mathbb{R}_+ \times E$ implies convergence in measure. The following argument is a minor adaptation of the proof of Lemma 1 of [18] which treats the case $E = \mathbb{R}$.

Classically, we can (homeomorphically) embed any Polish space E in the unit ball of a separable Hilbert space. For instance, E can be embed in the Hilbert cube $\phi : E \rightarrow [0, 1]^{\mathbb{N}}$ by considering the countable family of bounded continuous test functions $\phi_k(x) = d(x, x_k)$ where $(x_k)_{k \in \mathbb{N}}$ is a dense subset of E , and $d \leq 1$ is a bounded distance metrizing E ; indeed, a sequence converges in E towards a given limit in E if and only if its images by ϕ_k converges for each $k \geq 0$. As a consequence, the weighted ℓ_w^2 Hilbert space defined by the norm

$$\tilde{d}(x, y)^2 := \|\phi(x) - \phi(y)\|_{\ell_w^2}^2 := \sum_{k \geq 0} 2^{-k-1} |\phi_k(x) - \phi_k(y)|^2$$

also metrizes the Polish space E although \tilde{d} may not be complete.

We can next consider the Hilbert space $L^2(e^{-t}dt, \ell_w^2)$ of measurable paths on \mathbb{R}_+ taking values in the separable Hilbert ℓ_w^2 . By a routine finite dimensional approximation argument (one can consider increasing finite-dimensional sub-spaces $\ell_d \subset \ell_w^2$ with $l_d \nearrow d$ ℓ_w^2), the space of bounded continuous functions from \mathbb{R}_+ to ℓ_w^2 is dense in $L^2(e^{-t}dt, \ell_w^2)$. As a consequence, a converging sequence of pseudo-paths $x^n \rightarrow x$ (as distributions of $\text{Pr}(E \times [0, +\infty[))$) also converges weakly (by definition, and since x^n and x lie in the unit ball of ℓ_w^2) in $L^2(e^{-t}dt, \ell_w^2)$, as can be seen by considering the convergence of the scalar product $\int_0^\infty \langle \phi(x_t^n), \varphi(t) \rangle e^{-t}dt$ for any continuous and bounded path φ with values in ℓ_w^2 . In the same way, convergence as pseudo-paths implies that the Hilbert norm in $L^2(e^{-t}dt, \ell_w^2)$ of x^n converges to the one of x . Hence, x^n converges strongly to x in $L^2(e^{-t}dt, \ell_w^2)$. Since the latter two lie in the unit ball, strong convergence in $L^2(e^{-t}dt, \ell_w^2)$ is equivalent to convergence in probability in ℓ_w^2 or equivalently in E since convergence in probability does not depend on the metric. \square

Now, one can consider *random* pseudo-paths defined as classes of equivalence of measurable maps $X : (\Omega, \mathbb{P}) \rightarrow L^0(e^{-t}dt, E)$ identified with almost sure equality. It is thus possible consider the set of random pseudo-paths as the space $L^0(\mathbb{P}, L^0(e^{-t}dt, E))$. The following lemma is somehow trivial, but nonetheless clarifying.

Lemma A.2. *The metric*

$$\mathbb{E} \int_0^\infty \min(d(X_t, Y_t), 1) e^{-t} dt$$

is an isometry between $L^0(\mathbb{P}, L^0(e^{-t}dt, E))$ and $L^0(\mathbb{P} \otimes e^{-t}dt, E)$, when we metrize $L^0(e^{-t}dt, E)$ with (A.2). In particular, any random pseudo-path can be represented by a $\mathbb{R}_+ \times \Omega$ -measurable process $(t, \omega) \mapsto X(t, \omega)$.

Proof. Since E is separable, there exists $C \subset E$ a countable dense subset of E . Random variables taking values in C are dense for convergence in probability. The isometry is trivial for such random variables by considering all maps $(t, \omega) \mapsto X(t, \omega) \in C$ which are automatically measurable. \square

A.2 Convergence of finite-dimensional distributions

Convergence in pseudo-path space of càdlàg processes does not imply convergence of finite dimensional distribution (as is already the case for the Skorokhod topology). However, up to extraction, one can obtain convergence of finite dimensional distributions for all times in a set of full Lebesgue measure. This was proved as Theorem 5 in [18] for general real values processes. We sketch the proof again in the specific càdlàg case but for Polish-valued processes.

Lemma A.3. *Let E be Polish and let X^n be a sequence of càdlàg processes converging in distribution, for the pseudo-path topology, towards a càdlàg X^∞ process. Then there exists a sub-sequence and a subset $J \subset \mathbb{R}_+$ of full Lebesgue measure for which all finite-dimensional distributions converge.*

Proof. By the Skorokhod representation theorem in Polish spaces, one can assume that \tilde{X}^n converges towards \tilde{X}^∞ almost surely, where \tilde{X}^n has the same pseudo-path distribution as X_n , for each $n \geq 1$. Since by assumption $\mathbb{P}(\tilde{X}_n \in \mathbb{D}(\mathbb{R}_+, E) \subset L^0(e^{-t}dt, E)) = 1$, we can canonically consider \tilde{X}^n as a càdlàg process using its unique càdlàg representative. We thus subsequently drop the $\tilde{}$ notation.

By Lemma A.1, almost sure convergence in pseudo-paths space implies almost sure convergence in probability. Taking the expectation and using dominated convergence one gets

$$\int_0^\infty \mathbb{E}[\min(d(X_t^n, X_t^\infty), 1)] e^{-t} dt \xrightarrow{n \rightarrow \infty} 0,$$

that is convergence in probability but with respect to the product measure $e^{-t}dt \otimes \mathbb{P}$. As a consequence there exists a sub-sequence (we do not change notation to denote this sub-sequence) and a set of full $e^{-t}dt \otimes \mathbb{P}$ measure on which X^n converges towards X^∞ . Using Fubini, this means that there is a set J of full Lebesgue measure for which $t \in J$ implies that the probability $\lim_n X_t^n = X^\infty$ is one. This implies convergence of finite dimensional distributions for all times in J . \square

A.3 Compact sub-spaces

Compact subsets of the pseudo-paths space can be easily characterized using a direct application of the Prokhorov theorem.

Lemma A.4. *A subset \mathcal{K} of pseudo-paths is relatively compact if and only if for any $\varepsilon > 0$, there exists $t_\varepsilon > 0$ and $K_\varepsilon \subset E$ compact such that*

$$\inf_{x \in \mathcal{K}} \int_0^{t_\varepsilon} \mathbb{1}_{d(x_t, K_\varepsilon) \leq \varepsilon} e^{-t} dt \geq 1 - \varepsilon;$$

in particular, if for all $T > 0$, all the pseudo-paths of \mathcal{K} are taking on $[0, T]$ their values in a compact $K_T \subset E$, then the pseudo-paths space \mathcal{K} is compact.

A.4 The Meyer-Zheng criterion

We now turn to the case where $E = \mathbb{R}$. The idea is to consider the concept of *mean variation* on pseudo-path distributions. A càdlàg random process with finite mean variations characterizes (integrable) càdlàg quasi-martingales.

Definition A.5 (Mean variation). *Let $T > 0$ denotes an horizon time, and $(X_t)_{t \geq 0}$ a càdlàg real valued random process with $Z_t \in L^1$ for all t . The mean variation of Z over $[0, T]$ with respect to its own natural filtration is given by*

$$\begin{aligned} V_T(Z) &:= \sup_{0=t_0 \leq t_1 \leq \dots \leq t_K \leq T} \sum_{k=0}^{K-1} \mathbb{E} \left| \mathbb{E} [Z_{t_{k+1}} - Z_{t_k} \mid \sigma(Z_t, t \leq t_k)] \right|, \\ &= \sup_{|H| \leq 1} \mathbb{E} \left(\int_0^T H_t dZ_t \right), \end{aligned}$$

where in the above the supremum is taken over predictable processes (with respect to the natural filtration of Z) taking value in $[-1, 1]$.

As noted in Remark 1 p. 362 of [18] the mean variations is increasing with the considered filtration, so that proving a uniform bound for a larger filtration immediately yields the same bound for the natural filtration of the considered process.

The main theorem of Meyer-Zheng (Theorem 4) can be presented as follows.

Theorem A.6 (Meyer-Zheng criteria). *Let $(\pi_n)_{n \geq 1}$ denotes a sequence of distributions of real valued càdlàg processes. Assume that this sequence is converging for the pseudo-path topology towards a distribution in pseudo-path space π . If for all $T > 0$ the mean variation is uniformly bounded $\limsup_n V_T(\pi_n) < +\infty$, then the pseudo-path limit is càdlàg, that is π_∞ fully charges $\mathbb{D}(\mathbb{R}_+, \mathbb{R}) \subset L^0(e^{-t}dt, \mathbb{R})$, and it has bounded mean variation: $V_T(\pi) \leq \limsup_n V_T(\pi_n) < +\infty$ for all T .*

Proof. This is a direct adaptation of Theorem 4 in [18].

Let $T > 0$ by given and consider the distribution $(\pi_n^T)_{n \geq 1}$ and π^T of pseudo-paths stopped at T (equal to their value at T for all $t \geq T$). The sequence $(\pi_n^T)_{n \geq 1}$ again converges towards π^T . Using the proof of Theorem 4 in [18] after the preliminary extraction step (which is superfluous here), we get that π^T has support in \mathbb{D} . Using Remark 2 after the proof, we also obtain that $V_T(\pi) \leq \limsup_n V_T(\pi_n) < +\infty$.

Finally, in order to remove T , one can naively remark that a pseudo-path x has a càdlàg representative if and only if the stopped pseudo-path x^T has a càdlàg representative for all integer T :

$$x \in \mathbb{D}(\mathbb{R}_+, E) \subset L^0(e^{-t}dt, E) \Leftrightarrow x^T \in \mathbb{D}(\mathbb{R}_+, E) \subset L^0(e^{-t}dt, E), \forall T \in \mathbb{N},$$

by identifying the unique càdlàg representative on each interval $[T, T+1]$. By σ -additivity, it follows that π fully charges \mathbb{D} . □

A.5 Characterization of càdlàg martingales

We will also use in this article the following characterization of the martingale property on càdlàg processes:

Theorem A.7. *An integrable càdlàg real valued process \mathcal{M} is a martingale for the natural filtration of a càdlàg process X if and only if the following holds: for all $0 \leq t_1 \leq \dots \leq t_k \leq t$ in some dense subset of \mathbb{R}_+ , all $\varphi_1, \dots, \varphi_k$ continuous bounded test functions, and all $k \geq 1$, one has:*

$$\mathbb{E}[(\mathcal{M}_t - \mathcal{M}_{t_k})\varphi_k(X_{t_k})\dots\varphi_1(X_{t_1})] = 0.$$

It is then possible to use this property to prove that the limit of a sequence of martingales is again a martingale. Typically, the considered sequence of martingales is constructed from a sequence of càdlàg Markov processes that converge to a càdlàg process for the pseudo-path topology; and the convergence of almost all finite dimensional distributions is sufficient to a martingale (closure property). This will be carried out in Step 3 of Section 3 while proving the main theorem.

Appendix B Some remarks on diffusions

The next lemma will be used in the examples of Section 4 in order to check that strong solutions of SDEs do remain in a given closed subset of \mathbb{R}^n .

Lemma B.1. *Let $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ be a strong solution of a SDE with Lipschitz coefficients on \mathbb{R}^n . Assume that there is a smooth function $\xi : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $E = \{x : \xi(x) \leq 0\}$ and verifying on $\mathbb{R}^n \setminus E$: i) $L\xi := b \cdot \nabla \xi + \frac{1}{2} \sigma \sigma^T : \nabla^2 \xi \leq 0$, as well as ii) $\nabla \xi \sigma = 0$. Then X_t , starting from E remains in E for all time, that is E is an absorbing set.*

Proof. Consider the semi-martingale $Z_t := \xi^+(X_t)$. By construction, the latter is predictable, so its local time at 0 vanishes, and the Itô-Tanaka formula implies

$$dZ_t = \mathbb{1}_{Z_t > 0} L\xi(X_t)dt \leq 0.$$

As a consequence, if $x \in E$ then $Z_t = 0$ for all times, that is to say $X_t \in E$ for all times. \square

The next lemma will be used in the examples of Section 4 in order to check that strong solutions of SDEs do satisfy Assumption 4 and Assumption 6.

Lemma B.2. *Assume the initial process is solution in \mathbb{R}^n of a Stochastic Differential Equation of the form*

$$dX_t^{(\gamma_1, \gamma_0)} = \sqrt{\gamma_1} \sigma_1(X_t^{(\gamma_1, \gamma_0)}) dW_t^{(1)} + \gamma_1 b_1(X_t^{(\gamma_1, \gamma_0)}) dt + \sqrt{\gamma_0} \sigma_0(X_t^{(\gamma_1, \gamma_0)}) dW_t^{(0)} + \gamma_0 b_0(X_t^{(\gamma_1, \gamma_0)}) dt,$$

with bounded, globally Lipschitz coefficients $(\sigma_1, b_1, \sigma_0, b_0)$. Then Assumption 4 and Assumption 6 are satisfied.

Proof. For technical reason will here consider a more general SDE than the one considered above: instead of taking the coefficient in front of σ_0 equal to $\sqrt{\gamma_0}$, we add a parameter γ'_0 and we denote $X^{\gamma_1, \gamma_0, \gamma'_0}$ the process where we replaced $\sqrt{\gamma_0}$ by γ'_0 . Let us now consider the diffusion process on the (slightly) extended state space $E \times [0, 1] \times [0, 1]$ defined by:

$$t \mapsto (X_t^{(1, E_t, F_t)}, E_t, F_t)$$

in which $E_t = E_0 = e$ and $F_t = F_0 = f$ are constant over time. The latter is also a strong solution of the SDE defined on $E \times [0, 1] \times [0, 1]$ with Lipschitz coefficients, since the coefficient $(x, e, f) \mapsto (e \times b_0(x) + f \times \sigma_0, 0, 0)$ is obviously Lipschitz since b_0 and σ_0 are bounded. On the other hand, we know (see [24, Theorem 13.1]) that a strong solution of a SDE with Lipschitz coefficients can be represented by a stochastic flow $(x, e, f, t) \mapsto X_t^{(1, e, f)}(x)$ which is almost surely continuous. Taking in addition $f = \sqrt{e}$, this continuity implies both Assumption 4 and Assumption 6 by a routine dominated convergence argument and the Heine-Cantor theorem. \square

Appendix C Calculus associated with the diffusion in a strip

Analysis of $\mathcal{P}\varphi$

$\mathcal{P}\varphi$ is given by the following explicit expressions of $\mathbb{P}_y(\tau \in dt; W_\tau = 1)$ and $\mathbb{P}_y(\tau \in dt; W_\tau = -1)$ based on theta functions (See [6, Part II, Chap. 1, Section 3.0.6 and App.2, Section 11]):

$$\mathbb{P}_y(\tau \in dt; W_\tau = 1) = \sum_{k=-\infty}^{+\infty} \frac{4k + 1 - y}{\sqrt{2\pi t^{3/2}}} e^{-(4k+1-y)^2/(2t)} dt, \tag{C.1}$$

and:

$$\mathbb{P}_y(\tau \in dt; W_\tau = -1) = \sum_{k=-\infty}^{+\infty} \frac{4k + 1 + y}{\sqrt{2\pi t^{3/2}}} e^{-(4k+1+y)^2/(2t)} dt. \tag{C.2}$$

We now give a detailed elementary analysis of these functions. Note that it is possible to double-check it by remarking that $\mathbb{P}_y(\tau \in dt; W_\tau = 1) = \frac{1}{4} \partial_y \theta(\frac{1-y}{4}, i\pi t/8)$ where θ is the standard Jacobi theta function, and by using standard results on the latter.

For the sake of simplicity, we will mainly consider (C.1), the computations with (C.2) being identical by symmetry. We claim that if $\varphi \in C_c^\infty(\mathbb{R})$, then $y \mapsto \int_0^\infty \varphi(t) \mathbb{P}_y(\tau \in dt)$

$dt; W_\tau = 1$) is smooth on $[-1, 1]$. The terms in (C.1) for $k \neq 0$ are easy to handle by routine dominated convergence. For $k = 0$, we remark that a difficulty arises for $y = 1$, but that the associated term is of the form $g(1 - y)$ where g is the following smooth function:

Lemma C.1. *Let φ be smooth with compact support. The function on \mathbb{R}_+ defined by*

$$g(u) := \int_0^\infty e^{-u^2/t} u t^{-\frac{3}{2}} \varphi(t) dt$$

is smooth on \mathbb{R}_+ .

Proof. By a change of scale $u^2/t = 1/t'$:

$$\begin{aligned} g(u) &= \int_0^\infty e^{-1/t'} u (u^2 t')^{-\frac{3}{2}} \varphi(u^2 t') u^2 dt' \\ &= \int_0^\infty e^{-1/t'} u^{1-3+2} t'^{-\frac{3}{2}} \varphi(u^2 t') dt' \\ &= \int_0^\infty e^{-1/t'} t'^{-\frac{3}{2}} \varphi(u^2 t') dt'. \end{aligned}$$

Using that the support of φ is compact, we have that the function is continuous on \mathbb{R}_+ and continuously differentiable on \mathbb{R}_+^* . We have after a change of scale $u^2/s = 1/t'$:

$$\begin{aligned} g'(u) &= \int_0^\infty e^{-1/t'} t'^{-\frac{3}{2}} \varphi'(u^2 t') 2t' u dt' \\ &= 2 \int_0^\infty e^{-u^2/s} \left(\frac{u^2}{s}\right)^{\frac{1}{2}} \varphi'(s) u \frac{1}{u^2} ds \\ &= 2 \int_0^\infty e^{-u^2/s} s^{-\frac{1}{2}} \varphi'(s) ds \\ &\xrightarrow{u \rightarrow 0^+} 2 \int_0^{+\infty} s^{-\frac{1}{2}} \varphi'(s) ds. \end{aligned}$$

We have hence that g is continuously differentiable on \mathbb{R}_+ . Deriving a second time we find:

$$\begin{aligned} g''(u) &= 2 \int_0^{+\infty} e^{-\frac{u^2}{s}} \left(-\frac{2u}{s}\right) s^{-\frac{1}{2}} \varphi'(s) ds \\ &= -4 \int_0^{+\infty} e^{-\frac{u^2}{s}} u s^{-\frac{3}{2}} \varphi'(s) ds \end{aligned}$$

The form of this function is exactly the one of g up to multiply by factor -4 and replace φ by φ' . Since we prove that g is continuously differentiable on \mathbb{R} , we have by an immediate induction that g is in fact infinitely differentiable on \mathbb{R}_+ . \square

We therefore have that $y \mapsto \mathcal{P}(y, z)$ is smooth on $[-1, 1]$ for all $z \in \mathbb{R}$. Since $\mathcal{P}\varphi$ is trivially smooth in z by regularity of φ , we have that $(y, z) \mapsto \mathcal{P}\varphi(y, z)$ is smooth on $E = [-1, 1] \times \mathbb{R}$.

The measures μ (resp. ν) are positive Levy of a subordinator (resp. positive finite)

In this section we will study μ and ν and then prove that we can take the limit under the integral when y approaches ± 1 in $\partial_y \mathcal{P}\varphi(y, z)$. In order to do that, we will first give a simple expression of $\partial_y \mathcal{P}\varphi(y, z)$:

Lemma C.2. For all $y \in]-1, 1[$ we have:

$$\partial_y \mathbb{P}_y(\tau \in dt, W_\tau = 1) = -\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) k^2 e^{-t \frac{\pi^2 k^2}{8}} dt, \tag{C.3}$$

and:

$$\partial_y \mathbb{P}_y(\tau \in dt, W_\tau = -1) = \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1+y}{2}\right) k^2 e^{-t \frac{\pi^2 k^2}{8}} dt. \tag{C.4}$$

Proof. We start with the first formula. Deriving (C.1) under y once we obtain:

$$\begin{aligned} & \frac{\partial}{\partial y} \mathbb{P}_y(\tau \in dt, W_\tau = 1) \\ &= \sum_{k=-\infty}^{+\infty} -\frac{1}{\sqrt{2\pi t^{\frac{3}{2}}}} e^{-\frac{(4k+1-y)^2}{2t}} + \frac{4k+1-y}{\sqrt{2\pi t^{\frac{3}{2}}}} e^{-\frac{(4k+1-y)^2}{2t}} \times \frac{2(4k+1-y)}{2t} dt \\ &= \frac{1}{\sqrt{2\pi t^{\frac{3}{2}}}} \sum_{k=-\infty}^{+\infty} e^{-\frac{(4k+1-y)^2}{2t}} \left(-1 + \frac{(4k+1-y)^2}{t}\right) dt \end{aligned} \tag{C.5}$$

Writing $f : t \mapsto e^{-\pi a(t + \frac{1-y}{4})^2}$ and $\mathcal{F}f : t \mapsto \int_{\mathbb{R}} e^{-2i\pi t u} f(u) du$, Poisson summation formula gives us:

$$\sum_{k=-\infty}^{+\infty} f(k) = \sum_{k=-\infty}^{+\infty} \mathcal{F}f(k)$$

But, for $g : t \mapsto e^{-\pi a t^2}$, we have $(\tau_{-\frac{1-y}{4}}) * g = f$ and $\mathcal{F}g(t) = a^{-\frac{1}{2}} e^{-\frac{\pi}{a} t^2}$, thus:

$$\begin{aligned} \mathcal{F}f(k) &= \mathcal{F}((\tau_{-\frac{1-y}{4}}) * g)(k) \\ &= e^{ik\pi \frac{1-y}{2}} \mathcal{F}g(k) \\ &= e^{ik\pi \frac{1-y}{2}} a^{-\frac{1}{2}} e^{-\frac{\pi}{a} k^2} \end{aligned}$$

Eventually, we have:

$$\begin{aligned} \sum_{k=-\infty}^{+\infty} e^{-\pi a(k + \frac{1-y}{4})^2} &= \sum_{k=-\infty}^{+\infty} e^{ik\pi \frac{1-y}{2}} a^{-\frac{1}{2}} e^{-\frac{\pi}{a} k^2} \\ &= \frac{1}{2} \sum_{k=-\infty}^{+\infty} 2 \cos\left(k\pi \frac{1-y}{2}\right) a^{-\frac{1}{2}} e^{-\frac{\pi}{a} k^2} \\ &= \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) a^{-\frac{1}{2}} e^{-\frac{\pi}{a} k^2} \end{aligned}$$

If we differentiate this equation under a we get:

$$\begin{aligned} & \sum_{k=-\infty}^{+\infty} -\pi \left(k + \frac{1-y}{4}\right)^2 e^{-\pi a(k + \frac{1-y}{4})^2} \\ &= -\frac{1}{2} a^{-\frac{3}{2}} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) e^{-\frac{\pi}{a} k^2} + a^{-\frac{1}{2}} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) e^{-\frac{\pi}{a} k^2} \frac{\pi k^2}{a^2} \end{aligned}$$

Taking $a = \frac{16}{2\pi t}$, and we have respectively:

$$\sum_{k=-\infty}^{+\infty} e^{-\frac{(4k+1-y)^2}{2t}} = \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) \frac{\sqrt{2\pi t}}{4} e^{-t\frac{\pi^2 k^2}{8}}, \tag{C.6}$$

and:

$$\begin{aligned} & \sum_{k=-\infty}^{+\infty} -\pi \left(k + \frac{1-y}{4}\right)^2 e^{-\frac{(4k+1-y)^2}{2t}} \\ &= -\frac{1}{2} \frac{\sqrt{2\pi t}^3}{64} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) e^{-t\frac{\pi^2 k^2}{8}} \\ & \quad + \frac{\sqrt{2\pi t}}{4} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) e^{-t\frac{\pi^2 k^2}{8}} \frac{\pi^3 t^2 k^2}{64} \end{aligned} \tag{C.7}$$

We now multiply (C.6) by $-\frac{1}{\sqrt{2\pi t}^{\frac{3}{2}}}$ and (C.7) by $-\frac{1}{\sqrt{2\pi t}^{\frac{3}{2}}} \times \frac{16}{\pi t}$ to get:

$$\begin{aligned} & \frac{1}{\sqrt{2\pi t}^{\frac{3}{2}}} \sum_{k=-\infty}^{+\infty} e^{-\frac{(4k+1-y)^2}{2t}} \left(-1 + \frac{(4k+1-y)^2}{t}\right) \\ &= -\frac{1}{4t} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) e^{-t\frac{\pi^2 k^2}{8}} + \frac{1}{4t} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) e^{-t\frac{\pi^2 k^2}{8}} \\ & \quad - \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) k^2 e^{-t\frac{\pi^2 k^2}{8}} \\ &= -\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} \cos\left(k\pi \frac{1-y}{2}\right) k^2 e^{-t\frac{\pi^2 k^2}{8}} \end{aligned}$$

Combining this with (C.5), we get exactly (C.3). The proof of the second formula follows exactly the same steps, except we consider $\tau_{-\frac{1+y}{4}}$ instead of $\tau_{-\frac{1-y}{4}}$. \square

Let us take the limit in (C.3) when y approaches 1:

$$\lim_{y \rightarrow 1} \frac{\partial}{\partial y} \mathbb{P}_y(\tau \in dt, W_\tau = 1) = -\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 e^{-t\frac{\pi^2 k^2}{8}} dt \tag{C.8}$$

and define $\mu(dt) := \left(\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 e^{-t\frac{\pi^2 k^2}{8}}\right) dt$ if $t > 0$, while $\mu(dt) = 0$ if $t \leq 0$. This coincides with the definition of μ given in the text and we have that μ is trivially a positive measure.

Let us show that μ is the Levy measure of a subordinator. We have that $\mu \geq 0$ and:

$$\begin{aligned}
 & \int_0^{+\infty} \min(1, t) \mu(dt) \\
 &= \int_0^1 t \left(\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 e^{-t \frac{\pi^2 k^2}{8}} \right) dt + \int_1^{+\infty} \left(\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 e^{-t \frac{\pi^2 k^2}{8}} \right) dt \\
 &= \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 \int_0^1 t e^{-t \frac{\pi^2 k^2}{8}} dt + \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 \int_1^{+\infty} e^{-t \frac{\pi^2 k^2}{8}} dt \\
 &= \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 \left(\left[-\frac{8}{\pi^2 k^2} t e^{-t \frac{\pi^2 k^2}{8}} \right]_0^1 + \frac{8}{\pi^2 k^2} \int_0^1 e^{-t \frac{\pi^2 k^2}{8}} dt \right) \\
 &\quad + \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 \left[-\frac{8}{\pi^2 k^2} e^{-t \frac{\pi^2 k^2}{8}} \right]_1^{+\infty} \\
 &= \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 \left(-\frac{8}{\pi^2 k^2} e^{-\frac{\pi^2 k^2}{8}} + \left(\frac{8}{\pi^2 k^2} \right)^2 (1 - e^{-\frac{\pi^2 k^2}{8}}) \right) \\
 &\quad + \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 \frac{8}{\pi^2 k^2} e^{-\frac{\pi^2 k^2}{8}} \\
 &< +\infty
 \end{aligned} \tag{C.9}$$

Let us now take the limit in (C.4) when y approaches 1:

$$\lim_{y \rightarrow 1} \frac{\partial}{\partial y} \mathbb{P}_y(\tau \in dt, W_\tau = -1) = \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} dt, \tag{C.10}$$

and define $\nu(dt) := \left(-\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} \right) dt$ if $t > 0$, while $\nu(dt) = 0$ if $t \leq 0$.

This coincides with the definition of ν given in the text and we will now prove that this measure is positive and finite of mass

$$\int_{-\infty}^{+\infty} \nu(dt) = 1/2.$$

To show that $\sum_{k=-\infty}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} \leq 0$ for all $t > 0$, we will consider two cases.

We first assume that $t \geq 1$. Then the sequence $\left(k^2 e^{-t \frac{\pi^2 k^2}{8}} \right)_{k \geq 1}$ is decreasing since the function $x \mapsto x^2 e^{-t \frac{\pi^2 x^2}{8}}$ has a negative derivative for $x^2 \geq \frac{8}{t\pi^2}$ and thus for $x \geq 1$ in our situation. Hence for $t \geq 1$ we have:

$$\sum_{k=-\infty}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} = 2 \sum_{k=1}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} \leq 0$$

by the criterion of convergence of alternate series.

We now write:

$$\begin{aligned} \sum_{k=-\infty}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} &= -\frac{8}{\pi^2} \frac{\partial}{\partial t} \left[\sum_{k=-\infty}^{+\infty} (-1)^k e^{-t \frac{\pi^2 k^2}{8}} \right] \\ &= -\frac{8}{\pi^2} \frac{\partial}{\partial t} \left[\sum_{k=-\infty}^{+\infty} e^{-i\pi k - t \frac{\pi^2 k^2}{8}} \right]. \end{aligned} \tag{C.11}$$

We will next prove, in order to conclude on the positivity of ν , that $t \mapsto \sum_{k=-\infty}^{+\infty} e^{-i\pi k - t \frac{\pi^2 k^2}{8}}$ is increasing for $t \in]0, 1[$. Writing $h : x \mapsto e^{-t \frac{\pi^2 x^2}{8}}$, we have that:

$$\mathcal{F}(e^{-i\pi x} h)(\xi) = (\tau_{-\frac{1}{2}}) * \mathcal{F}(h)(\xi) = \sqrt{\frac{8}{\pi t}} e^{-\frac{s(\xi + \frac{1}{2})^2}{t}}$$

Thus, the Poisson summation formula gives:

$$F(t) := \sum_{k=-\infty}^{+\infty} e^{-i\pi k - t \frac{\pi^2 k^2}{8}} = \sum_{k=-\infty}^{+\infty} e^{-i\pi k} h(k) \tag{C.12}$$

$$= \sum_{k=-\infty}^{+\infty} \mathcal{F}(e^{-i\pi x} h)(k) \tag{C.13}$$

$$= \sqrt{\frac{8}{\pi t}} \sum_{k=-\infty}^{+\infty} e^{-\frac{s(k + \frac{1}{2})^2}{t}} \tag{C.14}$$

We consider for any $k \in \mathbb{Z}$, the function $f_k : t \mapsto t^{-\frac{1}{2}} e^{-\frac{s(k + \frac{1}{2})^2}{t}}$, we have that $f'_k : t \mapsto e^{-\frac{s(k + \frac{1}{2})^2}{t}} t^{-\frac{3}{2}} \left[-\frac{1}{2} + \frac{8(k + \frac{1}{2})^2}{t} \right]$. Thus, $f'_k(t) \geq 0$ if and only if $-\frac{1}{2} + \frac{8(k + \frac{1}{2})^2}{t} \geq 0$, that is to say $t \leq 16(k + \frac{1}{2})^2$, but this inequality is true for every k if we have that $t \leq 4$. Since we assume that $t \in]0, 1[$, we have that F is increasing for $t \in]0, 1[$.

Hence one obtains finally, for all $t > 0$, $-\frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} (-1)^k k^2 e^{-t \frac{\pi^2 k^2}{8}} \geq 0$, and thus $\nu \geq 0$.

To compute the total mass of ν , consider (C.11) to get

$$\int_0^{+\infty} \nu(dt) = \frac{1}{2} (F(+\infty) - F(0)),$$

and remark that $F(+\infty) = 1$, while by (C.14) $F(0) = 0$.

Remark C.3. We also have that $\mu(dt) = \lim_{y \rightarrow -1} \partial_y \mathbb{P}_y(\tau \in dt, W_\tau = -1)$ and $\nu(dt) = \lim_{y \rightarrow -1} \partial_y \mathbb{P}_y(\tau \in dt, W_\tau = 1)$.

We proved the wanted properties on μ and ν , however we still have to show that we can take the limit under the integral in the explicit expression of $\partial_y \mathcal{P}(\varphi - \varphi(1, z))(y, z)$ to get (4.1). The key idea is to remark that, since $|\cos(k\pi \frac{1-y}{2})| \leq 1$ and $|\cos(k\pi \frac{1+y}{2})| \leq 1$, equations (C.3) and (C.4) give us that for all $y \in]-1, 1[$ and $t > 0$:

$$\max(|\partial_y \mathbb{P}_y(\tau \in dt, W_\tau = 1)|, |\partial_y \mathbb{P}_y(\tau \in dt, W_\tau = -1)|) \leq \frac{\pi^2}{16} \sum_{k=-\infty}^{+\infty} k^2 e^{-t \frac{\pi^2 k^2}{8}} dt.$$

Indeed, we proved in (C.9) that this quantity is integrable on \mathbb{R}_+ against $\min(1, t)$, and $\varphi(\pm 1, z + \cdot) - \varphi(\pm 1, z)$ is bounded by a constant time $\min(1, t)$, so by the dominated convergence theorem we can take the limit when y approaches ± 1 to get what we want.

Appendix D Calculus associated with the diffusion in the Euclidean ball

Elliptic regularity immediately gives us that if φ is smooth on the unit sphere, then $x \mapsto \mathcal{P}\varphi(x)$ is also smooth (see [12, Ch. 6]). For the sake of completeness we nonetheless give explicit calculations that are the most relevant to our context.

Continuity of the projector

Let us consider $x_k \in \overline{\mathbb{B}^n}$ that converges to x . The explicit formula for \mathcal{P} implies that $\mathcal{P}(x_k, \cdot) \rightarrow_k \mathcal{P}(x, \cdot)$ if either $x \in \mathbb{B}^n$ or if $x_k \in \mathbb{S}^{n-1}$ for all k , we can thus assume without loss of generality that $x_k \in \overline{\mathbb{B}^n}$ and $x \in \mathbb{S}^{n-1}$. Then we have for all $\alpha > 0$ and n such that $\|x - x_k\| \leq \frac{\alpha}{2}$:

$$\begin{aligned} \int_{\mathbb{S}^{n-1} \setminus B(x, \alpha)} \frac{1 - \|x_k\|^2}{\|x_k - y\|^n} d\sigma^{n-1}(y) &\leq (1 - \|x_k\|^2) \times \sup_{y \in \mathbb{S}^{n-1} \setminus B(x, \alpha)} \|x_k - y\|^{-n} \\ &\leq (1 - \|x_k\|^2) \left(\frac{2}{\alpha}\right)^n \\ &\xrightarrow[k \rightarrow +\infty]{} 0 \end{aligned} \tag{D.1}$$

Then, for any φ that is 1-Lipschitz on the ball, we have for all $\alpha > 0$ that

$$\begin{aligned} &\left| \int_{\mathbb{S}^{n-1}} \varphi(y) \frac{1 - \|x_k\|^2}{\|x_k - y\|^n} d\sigma^{n-1}(y) - \varphi(x) \right| \\ &= \left| \int_{\mathbb{S}^{n-1}} (\varphi(y) - \varphi(x)) \frac{1 - \|x_k\|^2}{\|x_k - y\|^n} d\sigma^{n-1}(y) \right| \\ &\leq \left| \int_{\mathbb{S}^{n-1} \setminus B(x, \alpha)} \|y - x\| \frac{1 - \|x_k\|^2}{\|x_k - y\|^n} d\sigma^{n-1}(y) \right| \\ &\quad + \left| \int_{B(x, \alpha)} \|y - x\| \frac{1 - \|x_k\|^2}{\|x_k - y\|^n} d\sigma^{n-1}(y) \right| \\ &\leq \left| \int_{\mathbb{S}^{n-1} \setminus B(x, \alpha)} 2 \times \frac{1 - \|x_k\|^2}{\|x_k - y\|^n} d\sigma^{n-1}(y) \right| + \left| \int_{B(x, \alpha)} \alpha \frac{1 - \|x_k\|^2}{\|x_k - y\|^n} d\sigma^{n-1}(y) \right|. \end{aligned}$$

The second term is inferior to α and using (D.1), we know that the first term converges to 0. Since this is true for any φ that is 1-Lipschitz, which are convergence determining, we have eventually that $\mathcal{P}(x_k, \cdot) \rightarrow_k \mathcal{P}(x, \cdot)$ in distribution.

Explicit expression of the radial derivative

For all $\varphi \in \mathcal{C}_c(\mathbb{S}^{n-1})$ and $a \in \mathbb{S}^{n-1}$, we write $f_a : \lambda \in]0, 1[\rightarrow \mathcal{P}\varphi(\lambda a)$. We have therefore for all $z \in \mathbb{B}^n \setminus \{0\}$, writing $z = \lambda a$ with $a \in \mathbb{S}^{n-1}$ and $\lambda = \|z\|$ that $\frac{\partial}{\partial r} \mathcal{P}\varphi(z) =$

$\frac{\partial}{\partial \lambda} f_a(\lambda)$. Let us compute this quantity:

$$\begin{aligned} & \frac{\partial}{\partial \lambda} f_a(\lambda) \\ &= \frac{\partial}{\partial \lambda} (f_a - \varphi(a))(\lambda) \\ &= \frac{\partial}{\partial \lambda} \int_{\mathbb{S}^{n-1}} \frac{1 - \lambda^2}{\|\lambda a - y\|^n} (\varphi(y) - \varphi(a)) \, d\sigma^{n-1}(y) \\ &= \frac{\partial}{\partial \lambda} \int_{\mathbb{S}^{n-1}} \frac{1 - \lambda^2}{\|\lambda a - y\|^n} (\varphi(y) - \varphi(a) - \langle T\varphi(a), y - a \rangle) \, d\sigma^{n-1}(y) \\ &\quad + \frac{\partial}{\partial \lambda} \int_{\mathbb{S}^{n-1}} \frac{1 - \lambda^2}{\|\lambda a - y\|^n} \langle T\varphi(a), y - a \rangle \, d\sigma^{n-1}(y) \end{aligned} \tag{D.2}$$

Let us denote by R_a the orthogonal reflection with unique invariant the axis generated by a . R_a is an isometry leaving invariant σ^{n-1} , $T\varphi(a) \perp a$ implies that $\langle T\varphi(a), R_a y - a \rangle = -\langle T\varphi(a), y - a \rangle$, and $\|\lambda a - R_a y\| = \|\lambda a - y\|$ (all these properties are trivial by symmetry). We conclude that the second term in (D.2) is the derivative of 0 and is hence null. Therefore:

$$\begin{aligned} & \frac{\partial}{\partial \lambda} f_a(\lambda) \\ &= \frac{\partial}{\partial \lambda} \int_{\mathbb{S}^{n-1}} \frac{1 - \lambda^2}{\|\lambda a - y\|^n} (\varphi(y) - \varphi(a) - \langle T\varphi(a), y - a \rangle) \, d\sigma^{n-1}(y) \\ &= \int_{\mathbb{S}^{n-1}} \frac{-2\lambda \|\lambda a - y\|^n - (1 - \lambda^2) n \frac{\langle \lambda a - y, a \rangle}{\|\lambda a - y\|} \|\lambda a - y\|^{n-1}}{\|\lambda a - y\|^{2n}} \\ &\quad \times (\varphi(y) - \varphi(a) - \langle T\varphi(a), y - a \rangle) \, d\sigma^{n-1}(y) \\ &= \int_{\mathbb{S}^{n-1}} \frac{-2\lambda - (1 - \lambda^2) n \frac{\langle \lambda a - y, a \rangle}{\|\lambda a - y\|} \|\lambda a - y\|^{-1}}{\|\lambda a - y\|^n} \\ &\quad \times (\varphi(y) - \varphi(a) - \langle T\varphi(a), y - a \rangle) \, d\sigma^{n-1}(y) \end{aligned}$$

Therefore, for $z \in \mathbb{B}^n$, we have:

$$\begin{aligned} \frac{\partial}{\partial r} \mathcal{P}\varphi(z) &= \int_{\mathbb{S}^{n-1}} \frac{-2\|z\| - (1 - \|z\|^2) n \frac{\langle z - y, a \rangle}{\|z - y\|} \|z - y\|^{-1}}{\|z - y\|^n} \\ &\quad \times \left(\varphi(y) - \varphi\left(\frac{z}{\|z\|}\right) - \langle T\varphi\left(\frac{z}{\|z\|}\right), y - \frac{z}{\|z\|} \rangle \right) \, d\sigma^{n-1}(y). \end{aligned}$$

Taking the limit when z approaches a (so in particular $\|z\|$ approaches 1) we find that:

$$\begin{aligned} & \lim_{z \rightarrow a} \frac{\partial}{\partial r} \mathcal{P}\varphi(z) \\ &= -2 \lim_{z \rightarrow a} \|z\| \int_{\mathbb{S}^{n-1}} \frac{1}{\|z - y\|^n} \left(\varphi(y) - \varphi\left(\frac{z}{\|z\|}\right) - \langle T\varphi\left(\frac{z}{\|z\|}\right), y - \frac{z}{\|z\|} \rangle \right) \, d\sigma^{n-1}(y) \\ &\quad - \lim_{z \rightarrow a} (1 - \|z\|^2) \int_{\mathbb{S}^{n-1}} \frac{n \frac{\langle z - y, a \rangle}{\|z - y\|} \|z - y\|^{-1}}{\|z - y\|^n} \\ &\quad \times \left(\varphi(y) - \varphi\left(\frac{z}{\|z\|}\right) - \langle T\varphi\left(\frac{z}{\|z\|}\right), y - \frac{z}{\|z\|} \rangle \right) \, d\sigma^{n-1}(y). \end{aligned} \tag{D.3}$$

For the first term in (D.3), we observe that the integrand converges simply to what we are looking for when z approaches a . We now denote R_z for $z \in \mathbb{B} \setminus \{0\}$ the only rotation

sending $\frac{z}{\|z\|}$ on a . Then, using that the measure σ is invariant by rotations and that rotations are isometries, we have that:

$$\begin{aligned} & \int_{\mathbb{S}^{n-1}} \frac{1}{\|z-y\|^n} \left(\varphi(y) - \varphi\left(\frac{z}{\|z\|}\right) - \left\langle T\varphi\left(\frac{z}{\|z\|}\right), y - \frac{z}{\|z\|} \right\rangle \right) d\sigma^{n-1}(y) \\ &= \int_{\mathbb{S}^{n-1}} \frac{1}{\|R_z(z) - y\|^n} \left(\varphi(R_z^{-1}(y)) - \varphi\left(\frac{z}{\|z\|}\right) - \left\langle T\varphi\left(\frac{z}{\|z\|}\right), R_z^{-1}(y) - \frac{z}{\|z\|} \right\rangle \right) d\sigma^{n-1}(y) \end{aligned} \tag{D.4}$$

Since φ in in $\mathcal{C}^2(\mathbb{S})$, there exists $M > 0$ such that:

$$\left| \left(\varphi(R_z^{-1}(y)) - \varphi\left(\frac{z}{\|z\|}\right) - \left\langle T\varphi\left(\frac{z}{\|z\|}\right), R_z^{-1}(y) - \frac{z}{\|z\|} \right\rangle \right) \right| \leq M \|y - a\|^2.$$

Furthermore $R_z(z) = \|z\|a$, so eventually the integrand in (D.4) is dominated by $2^n M \|y - a\|^{-(n-2)}$ that is integrable on the sphere, and we can use the dominated convergence theorem to conclude.

Let us now prove that the second term in (D.3) is equal to 0. We notice that $\left| \frac{\langle z-y, a \rangle}{\|z-y\|} \right| \leq 1$ and that $\varphi(y) - \varphi\left(\frac{z}{\|z\|}\right) - \left\langle T\varphi\left(\frac{z}{\|z\|}\right), y - \frac{z}{\|z\|} \right\rangle$ is dominated by $4M \|z - y\|^2$, so eventually the term inside the integral is dominated by $4nM \|z - y\|^{-(n-1)}$. Using a polar change of coordinates, we get that the integral on the sphere is hence dominated by $4nM |\ln(\|z - \frac{z}{\|z\|}\|)| = 4nM |\ln(1 - \|z\|)|$, but

$$(1 - \|z\|^2) |\ln(1 - \|z\|)| \xrightarrow{z \rightarrow a} 0,$$

so we have (4.4).

Acknowledgements

The authors want to thank Raphaël Chétrite and Cédric Bernardin for their help and support. This work has been partially supported by ANR SINEQ, ANR-21-CE40-0006.

References

- [1] Alberto Barchielli and Matteo Gregoratti. *Quantum trajectories and measurements in continuous time: the diffusive case*. Lect. Notes Phys. 782. Springer, 2009.
- [2] Tristan Benoist, Cédric Bernardin, Raphaël Chétrite, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. Emergence of jumps in quantum trajectories via homogenization. *Communications in Mathematical Physics*, 387(3):1821–1867, 2021.
- [3] Cédric Bernardin, Raphaël Chétrite, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. Spiking and collapsing in large noise limits of sdes. *The Annals of Applied Probability*, 33(1):417–446, 2023.
- [4] Cédric Bernardin, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. To spike or not to spike: the whims of the Wonham filter in the strong noise regime. *arXiv:2211.02032*, 2022.
- [5] Vladimir I. Bogachev. *Measure Theory*, volume 2. Springer Berlin Heidelberg, 2007.
- [6] Andrei N. Borodin and Paavo Salminen. *Handbook of Brownian Motion - Facts and Formulae*. Birkhäuser, 2002.
- [7] Heinz-Peter Breuer and Francesco Petruccione. *The theory of open quantum systems*. Oxford University Press, 2002.
- [8] Claude Dellacherie and Paul André Meyer. *Probabilités et potentiel: Chap. I à IV*. Hermann, 1975.
- [9] Joseph L. Doob. *Classical Potential Theory and Its Probabilistic Counterpart*. Classics in Mathematics. Springer, 1984.

- [10] Richard M. Dudley. *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2002.
- [11] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons Inc, 1986.
- [12] Lawrence C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.
- [13] Dimitri Faure. Averaging of semigroups associated to diffusion processes on a simplex. *Stochastic Processes and their Applications*, 150:323–357, 2022.
- [14] Serge Haroche and Jean-Michel Raimond. *Exploring the quantum: atoms, cavities, and photons*. Oxford university press, 2006.
- [15] Pei Hsu. On excursions of reflecting brownian motion. *Transactions of the American Mathematical Society*, 296(1):239–264, 1986.
- [16] Alexander S. Kechris. *Classical Descriptive Set Theory*. Graduate Texts in Mathematics. Springer New York, 2012.
- [17] Thomas G. Kurtz. Equivalence of stochastic equations and martingale problems. In *Stochastic analysis 2010*, pages 113–130. Springer, 2011.
- [18] Paul-André Meyer and Wei-An Zheng. Tightness criteria for laws of semimartingales. *Annales de l’IHP Probabilités et statistiques*, 20(4):353–372, 1984.
- [19] Gerard J. Milburn and Howard M. Wiseman. *Quantum Measurement and Control*. Cambridge University Press, 2010.
- [20] George Papanicolaou. Martingale approach to some limit theorems. In *Papers from the Duke Turbulence Conference, Duke Univ., Durham, NC, 1977, 1977*.
- [21] Étienne Pardoux. Homogenization of linear and semilinear second order parabolic pdes with periodic coefficients: a probabilistic approach. *Journal of Functional Analysis*, 167(2):498–520, 1999.
- [22] Étienne Pardoux and Alexander Yu. Veretennikov. Averaging of backward stochastic differential equations, with application to semi-linear pde’s. *Stochastics and Stochastic Reports*, 60(3-4):255–270, 1997.
- [23] Andrey Pilipenko. *An introduction to stochastic differential equations with reflection*. Lectures in Pure and Applied Mathematics 1. Universitätsverlag Potsdam, 2014.
- [24] L. Chris G. Rogers and David Williams. *Diffusions, Markov Processes, and Martingales*, volume 2. Cambridge University Press, 1979.
- [25] John Sylvester and Gunther Uhlmann. A global uniqueness theorem for an inverse boundary value problem. *Annals of Mathematics*, pages 153–169, 1987.
- [26] Antoine Tilloy, Michel Bauer, and Denis Bernard. Spikes in quantum trajectories. *Physical Review A*, 92(5):052111, 2015.