Long time behaviour of infinite dimensional stochastic processes

a thesis presented for the degree of

Doctor of Philosophy of Imperial College London

and the

Diploma of Imperial College

by

František Žák

Department of Mathematics

Imperial College

180 Queen's Gate, London SW7 2BZ

September 2016

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed: ______________________________

Copyright

Thesis advisor: Professor Boguslaw Zegarliński František Žák

Long time behaviour of infinite dimensional stochastic processes

Abstract

We study two examples of infinite dimensional stochastic processes. Situations and techniques involved are quite varied, however in both cases we achieve a progress in describing their long time behaviour.

The first case concerns interacting particle system of diffusions. We construct rigorously the process using finite dimensional approximation and the notion of martingale solution. The existence of invariant measure for the process is proved. The novelty of the results lies in the fact, that our methods enable us to consider such examples, where the generator of the diffusion is subelliptic.

The other project is related to stochastic partial differential equations and their stability properties. In particular it is shown that Robbins-Monro procedure can be extended to infinite dimensional setting. Thus we achieve results about pathwise convergence of solution. To be able to define corresponding solution, we rely on so-called variational approach to stochastic partial differential equations as pioneered by E. Pardoux, N. Krylov and B. Rozovskii. Our examples covers situations such as p-Laplace operator or Porous medium operator.

To my family, girlfriend Olga Globina, first teacher Libuše Fialová and first scientific advisor Josef Štěpán.

Acknowledgments

I would like to express my gratitude towards my supervisor prof Boguslaw Zegarliński, who helped me to settle in London, provided me with guidance through the whole PhD program. He always held belief, that I would be able to successfully finish the studies and gave me support and encouragement in all situations.

I would not be able to complete my studies without enormous help from Dr Jan Seidler, who was initially willing to consult the scientific issues in his spare time with me, and later on became my scientific collaborator for the second project of this thesis.

I wish to mention also my gratitude towards Dr Laurent Miclo and Dr Jan Swart, who gave me valuable feedback and offer me important ideas, how to tackle some of the challenging problems related to interacting particle system project. Dr Martin Ondreját provided very useful discussion concerning the stochastic approximation issues.

Immense acknowledgement I owe to my previous teachers in Prague, who showed me the beauty of abstract mathematics and whose influence led me to consider undertaking PhD. Most notably I am thankful to prof Luděk Zajíček for his great Analysis classes, prof Bohdan Maslowski for mentorship and believing in me, and to prof Josef Štěpán, who is unfortunately no longer amongst us.

Contents

List of Publications

The research presented in the chapters two and three has been made into an article, which has been recently published in journal -

- Frantisek Zak, *Existence of equilibrium for infinite System of interacting diffusions*, Stochastic Anal. and Appl. **34** (2016), 1057–1082,[83]
  Abstract: "We develop and implement new probabilistic strategy for proving basic results about long time behaviour for interacting diffusion processes on unbounded lattice. The concept of the solution used is rather weak as we construct the process as a solution to suitable infinite dimensional martingale problem. However the techniques allow us to consider cases where the generator of the particle is degenerate elliptic operator. As a model example we present situation, where the operator arises from Heisenberg group. In the last section we mention some further examples that can be handled using our methods. "

Chapter five is based upon another article, that is currently in review process:

- Jan Seidler, Frantisek Zak, *A note on Continuous-time stochastic approximations in Infinite Dimensions*,[73]
  Abstract: " We find sufficient conditions for convergence of a continuous-time Robbins-Monro type stochastic approximation procedure in infinite-dimensional Hilbert spaces in terms of Lyapunov functions, the variational approach to nonlinear stochastic partial differential equations being used as the main tool. "

People who know math understand what other mortals understand, but other mortals do not understand them. This asymmetry gives them a presumption of superior ability.

Daniel Kahneman

0

Introduction

In the given thesis we are concerned with stochastic processes and are particularly interested in describing their long time behaviour. From the physical point of view, the processes we study really are d-dimensional. Nevertheless, as it has been shown in the last thirty years or so, the use of infinite dimensional spaces to study these kind of processes often leads to important progress, as one can employ many useful tools that are available. Hence in the mathematical sense it is appropriate to describe the processes we study as infinite dimensional.

The thesis basically consist of two parts, that are connected by common techniques, interests and notions. We begin with the study of interacting particle system of diffusions. That is, we analyse d-dimensional lattice $\mathbb{Z}^d$, where we put a copy of diffusion particle living in $\mathbb{R}^n$ on each site and add finite range interactions between the particles. Original results are presented in chapters 2 and 3. At first we use now standard theory of Meyn and Tweedie to establish ergodicity results for finite lattice in chapter 2. The construction of infinite dimensional process and its invariant measure is presented in chapter 3. Our approach to the construction using solution to the martingale problem seems new in this context.

Second project is devoted to the extension of continuous time Robinson-Monro stochastic approximation to infinite dimension. Therefore here we establish path-wise convergence results about given stochastic PDEs, to which the theory can be applied. This work has been done by the author in conjunction with Dr Jan Seidler. Because we use the relatively heavy machinery of variational approach to stochastic equations in infinite dimension, we provide some necessary background in chapter 4. Chapter 5 then includes the general result about stochastic approximation in infinite dimension. We also list three particular examples of stochastic PDEs, that can be covered by this general theorem.

In chapter 1 we summarize some classical results about long time behaviour of stochastic processes that are useful in our work.

In conclusion in chapter 6 we list questions that arise from the thesis and remain unsolved.

0.1 Notation

Here we give general summary of the notations used throughout the thesis, for the comfort of reader we may occasionally repeat it, when used precisely. The related definitions of more advanced notions is provided in the work on appropriate place. The explanation for some less prevalent notation is also given in the subsequent text

$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$	natural numbers, integers, rationals, real numbers
$\mathbb{R}^\infty$	the space of all real valued infinite sequences
$a \wedge b$	minimum of a and$b$
$a \vee b$	maximum of a and b
$A \lesssim B$	$\Leftrightarrow \exists C > 0 : \; A < BC$
I_A	indicator (characteristic) function of set A
A^C	complement of the set A
∂A	boundary of the set A
$\|A\|$	cardinality of the set A
$A \subset\subset \mathbb{Z}^d$	subset with finite cardinality, i. e. $\|A\| < \infty$
clo A	closure of the set A
f^+	positive part of function f, i. e. $f^+ = \max(f, 0)$
f^-	negative part of function f, i. e. $f^- = \max(-f, 0)$
$o(g)$	function $f = o(g)$ at $a \in mathbb\mathbb{R}$, if $\lim_{x \to a} \frac{f(x)}{g(x)} = 0$
$i \in \mathbb{Z}^d : \\|i\\|_{\max}$	$\\|i\\|_{\max} = \max_{1 \le j \le d} \|i_j\|$
$x \in \mathbb{R}^n : \\|x\\|_{\mathbb{R}^n} = \\|x\\|_n = \\|x\\|$	Euclid norm of vector in $\mathbb{R}^n$
$\pi_k : \mathbb{R}^\infty \to \mathbb{R}^k$	projection function into first k coordinates, i. e. $\pi_k(x_1, x_2, \ldots) = (x_1, \ldots, x_k)$
$(a_x, a_y, a_z) \in \mathbb{R}^3 : \\|a\\|_{\mathbb{H}}$	$\\|a\\|_{\mathbb{H}} = \sqrt[4]{(a_x^2 + a_y^2)^2 + a_z^2}$
λ	Lebesgue measure
$\mu \circ f^{-1}$	the image of measure μ under mapping f
$\xrightarrow{w}$	weak convergence of measures

For a given metric spaces X, Y and $f : X \to Y$ we denote

$C(X, Y)$	space of continuous functions

If $Y = \mathbb{R}$ we simply write $C(X)$ and similarly for other type of spaces.

$C_b(X)$	space of bounded continuous functions
$f \in C_b(X) : \|f\|_\infty$	$\|f\|_\infty = \sup_{x \in X} \lvert f(x) \rvert$
$C_c(X)$	space of continuous functions with compact support
$C^p(X)$	space of p-times continuously differentiable functions (usually either $p = 2$ or $p = \infty$)
$\Delta f, f \in C^2(\mathbb{R}^d)$	Laplace operator on $\mathbb{R}^d$, i. e. $\Delta f = \sum_{i=1}^d \frac{\partial^2 f}{\partial^2 x_i}$
$\nabla f, f \in C^1(\mathbb{R}^d)$	gradient of function f in $\mathbb{R}^d$, i. e. $\nabla f = (\frac{\partial f}{\partial x_1}, \cdot, \frac{\partial f}{\partial x_d})$
$\mathcal{B}(X)$	space of Borel measurable functions

If $X \subset \mathbb{R}^\infty$, dim $X = \infty$ we have the notion of $f : X \to \mathbb{R}$ cylindrical function

$C^{Cyl}(X)$	space of cylindrical functions
$C_c^{Cyl}(X)$	space of cylindrical compactly supported functions

For Banach spaces E, F we denote

I	Identity operator on E, i. e. $Ix = x, x \in E$
E^*	dual space to the space E
${}_{E^*}\langle f, e \rangle_E$	dual pairing between E^* and E, i. e.

$$ {}_{E^*}\langle f, e \rangle_E = f(e), \;\; f \in E^*, e \in E $$

$\mathscr{L}(E, F)$	space of bounded linear operators from E to F
$L \in \mathscr{L}(E, F)$, L^*	adjoint operator

If both E and F are Hilbert spaces

$\langle \cdot, \cdot \rangle_E$	inner product on Hilbert space E
$\mathscr{L}_2(E, F)$	ideal of Hilbert-Schmidt operators in $\mathscr{L}(E, F)$

For open set $G \subseteq E$ and $f : G \to \mathbb{R}$

$Df(x)$	Gâteaux derivative of f at point x
$D^2f(x)$	second Gâteaux derivative of f at point x

For $O \subseteq \mathbb{R}^n$ open we use standard notation

$W^{k,p}(O)$	Sobolev space of functions with weak derivatives up to order k having finite L^p norm

1

Limit behaviour of stochastic processes

The concept of stochastic process is very broad and it is impossible to prove interesting results about their limit behaviour at this level of generality. Two important classes that enable non-trivial results are Markov processes and Martingales. We outline the basic theory for these classes, as our results in chapters 2 and 5 heavily rely on these classical results.

1.1 Ergodic theory of Markov processes

Our exposition in this section follows mostly lecture notes by M. Hairer[32], L. Rey-Bellet[71] and first part of the book[22] by G. Da Prato and J. Zabczyk. Intuitively Markov processes are processes without memory, its future movement depends only on current state. The modern view on Markov process has been established following foundational work of Dynkin and his students in 1950s, published in the monograph in 1959[27]. Thorough modern account of the general theory can be found in books[26], [12]. Since then Markov process is defined as a family of processes tied by the transition function, rather than a single process.

Definition 1.1.1 ((homogeneous) Markov process)**.** Let $\mathcal{S}$ be separable metric space equipped with standard Borel σ-field. A Markov process (X_t, P^x) is a stochastic process

$$X : [0, \infty) \times \Omega \to \mathcal{S}$$

and set of probability measures $\{P^x : x \in \mathcal{S}\}$ on $(\Omega, \mathcal{F})$ satisfying the following:

1) For each t, X_t is $\mathcal{F}_t$ measurable

2) For each t and each Borel subset A of $\mathcal{S}$, the map $x \to P^x(X_t \in A)$ is Borel measurable

3) For each $s, t \geq 0$, each Borel subset A of $\mathcal{S}$, and each $x \in \mathcal{S}$, we have

$$P^x(X_{s+t} \in A|\mathcal{F}_s) = \phi(X_s), \;\; \phi(\cdot) = P^{\cdot}(X_t \in A).$$

With every Markov process one naturally associates so called transition probability function $P_t : \mathcal{S} \times \mathcal{B}(\mathcal{S}) \to [0, 1]$

$$P_t(x, A) = P^x(X_t \in A), \;\; x \in \mathcal{S}, A \in \mathcal{B}_b(\mathcal{S}). \tag{1.1}$$

Easy computation using point 3) in definition of Markov process (denoting by E^x the expectation with respect to P^x) reveals that

$$P_{t+s}(x, A) = E^x P^x(X_{t+s} \in A|\mathcal{F}_s) = E^x P^{X_s}(X_t \in A) = \int_{\mathcal{S}} P_t(y, A) P_s(x, dy). \tag{1.2}$$

This relationship is called Chapman-Kolmogorov equation. In another words if we view P_t as an operator on bounded Borel functions on by putting

$$P_t f(\cdot) = \int_{\mathcal{S}} f(X_t) dP^{\cdot}, f \in \mathcal{B}(\mathcal{S})$$

then passing from indicator functions in (1.2) to general functions shows that P_t satisfies the semigroup property $P_{t+s} = P_t \circ P_s$.

Dual to the semigroup P_t we define P_t^* the semigroup acting on probability measures on $\mathcal{S}$ as

$$P_t^*(\mu)(\cdot) = \int_{\mathcal{S}} P_t(x, \cdot)d\mu(x).$$

Definition 1.1.2 (Invariant measure)**.** Measure μ is said to be invariant measure for the Markov process (X, P^x), if it is fixed point of semigroup P_t^*, i. e. $P_t^*\mu = \mu$ for all $t \geq 0$.

This definition is natural, because the process has stationary distribution if μ is its initial distribution. To see this, denote

$$P^\mu(X_t \in A) = \int_{\mathcal{S}} P^x(X_t \in A)d\mu(x) = \int_{\mathcal{S}} P_t(x, A)d\mu(x)$$

the distribution of process starting from μ. Realize that for times $t_1 < t_2$ and $A_1, A_2 \in \mathcal{B}(\mathcal{S})$ Markov property implies

$$P^x(X_{t_1} \in A_1, X_{t_2} \in A_2) = E^x P^x(X_{t_1} \in A_1, X_{t_2} \in A_2|\mathcal{F}_{t_1}) =$$

$$E^x I_{[X_{t_1} \in A_1]} P^x(X_{t_2} \in A_2|\mathcal{F}_{t_1}) = E^x I_{[X_{t_1} \in A_1]} P_{t_2 - t_1}(X_{t_1}, A_2) =$$

$$\int_{A_1} P_{t_2-t_1}(y, A_2)(P^x \circ X_{t_1}^{-1})(dy) = \int_{A_1} P_{t_2-t_1}(y, A_2)P_{t_1}(x, dy).$$

Hence if $P_t^*\mu = \mu$, by induction from previous we get for times $t_1 < \cdots < t_n$

$$P^\mu(X_{t_1+h} \in A_1, \ldots, X_{t_n+h} \in A_n)$$

$$= \int_{\mathcal{S}} \int_{A_1} \cdots \int_{A_{n-1}} P_{t_n - t_{n-1}}(y_{n-1}, A_n) \cdots P_{t_1+h}(x, dy_1)d\mu(x) =$$

$$\int_{\mathcal{S}} \int_{A_1} \cdots \int_{A_{n-1}} P_{t_n - t_{n-1}}(y_{n-1}, A_n) \cdots P_{t_1}(x, dy_1)d\mu(x) = P^\mu(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n),$$

and X is a stationary process as claimed.
It suffices to find measure μ satisfying $P_T^*\mu = \mu$ for some $T > 0$, as the measure $\hat{\mu}$

$$\hat{\mu}(A) = \frac{1}{T} \int_0^T P_t^*\mu(A)dt$$

is then invariant for the semigroup P_t. Indeed, we for any $f \in C_b(\mathcal{S})$ we compute

$$\hat{\mu} f = \frac{1}{T}\int_0^T (P_t^*\mu) f dt$$

$$P_s^*\hat{\mu} f = \int_{\mathcal{S}} P_s f d\hat{\mu} = \frac{1}{T}\int_0^T (P_t^*\mu) P_s f dt = \frac{1}{T}\int_0^T (P_{t+s}^*\mu) f dt =$$
$$\frac{1}{T}\left(\int_0^T (P_t^*\mu) f dt + \int_T^{T+s} (P_t^*\mu) f dt - \int_0^s (P_t^*\mu) f dt\right),$$

and the last the integrals cancel, since we assume $P_T^*\mu = \mu$. Therefore in the study of the questions of existence and uniqueness of invariant measure one can in principle focus on discrete time only.

Definition 1.1.3 (Feller properties)**.** We say Markov semigroup is Feller, if it maps continuous bounded functions in itself, i. e.

$$P_t f \in C_b, \ \forall f \in C_b(\mathcal{S}).$$

P_t is strong Feller, if it has smoothing effect, i. e.

$$P_t f \in C_b, \ \forall \mathcal{B}_b(\mathcal{S}).$$

The basic method for finding invariant measure dates back to 1937 and famous work of Krylov nad Bogoliubov[13] on dynamical systems. The method is robust enough that is easily adapted to the context of Markov processes.

Theorem 1.1.1 (Krylov-Bogoliubov)**.** *Let (X, P^x) be a Markov process on separable metric space $\mathcal{S}$ with transition function P_t. If the semigroup is Feller and there exists point $x \in \mathcal{S}$ such that the set of measures*

$$\mathcal{M} = \left\{\frac{1}{T}\int_0^T P_t(x,\cdot)dt\right\} \tag{1.3}$$

is tight, then there exists invariant measure for the process.

In fact the condition (1.3) is basically necessary for Feller semigroups, see[38] pp. 65.

Proof. By Prokhorov's theorem let μ be such measure that

$$\frac{1}{T_n}\int_0^{T_n} P_t(x,\cdot)dt \xrightarrow{w} \mu.$$

With the use of Feller property for arbitrary $f \in C_b$ one computes

$$\mu f = \lim_{n\to\infty} \frac{1}{T_n}\int_0^{T_n} P_t f(x)dt$$

$$(P_s^*\mu)f = \lim_{n\to\infty} \frac{1}{T_n}\int_0^{T_n} P_{t+s} f(x)dt =$$

$$\lim_{n\to\infty} \frac{1}{T_n}\int_0^{T_n} P_t f(x)dt + \lim_{n\to\infty} \frac{1}{T_n}\int_{T_n}^{T_n+s} P_t f(x)dt - \lim_{n\to\infty} \frac{1}{T_n}\int_0^{s} P_t f(x)dt.$$

Consider that $\|P_t f\| \le \|f\| < \infty$, hence the last two terms are zero and we have the desired equality $(P_s^*\mu)f = \mu f,\ \forall f \in C_b(\mathcal{S})$. $\square$

Very mild condition for Markov semigroup is that of stochastic continuity ([22] pp. 12), that is the equality

$$\lim_{t\to 0} P_t f(x) = f(x),\ \forall f \in C_b(\mathcal{S}).$$

To tie precisely long time behaviour of the process with invariant measure, it is useful to recall the notion of ergodicity. We say invariant measure μ is ergodic, if any set A satisfying

$$P_t I_A = I_A \ \mu\text{-a.s. } \forall t > 0$$

is trivial, i. e. $\mu(A) \in \{0,1\}$. The basic theorem about ergodic processes, that is reformulation of famous Birkhoff pointwise ergodic theorem, we can state in our context as follows.

Theorem 1.1.2. *Let P_t be stochastically continuous Markov semigroup with invariant measure μ. Then the following is equivalent*

(i) μ is ergodic

(ii) For any $A, B \in \mathcal{B}(\mathcal{S})$

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \int_{\mathcal{S}} P_t I_A(x) I_B(x) d\mu(x) dt = \mu(A)\mu(B).$$

(iii) For arbitrary $f \in L^2(\mathcal{S}, \mu)$

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T P_t f(x) dt = \int_{\mathcal{S}} f(x) d\mu(x) \ \mu\text{-a.s.}.$$

Proof. See[22] pp. 26. □

For the measure μ we define its support as the smallest closed set of full measure, i. e.

$$\text{supp } \mu = \text{clo } (\cap A; \mu(A) = 1).$$

As a consequence of Theorem 1.1.2 we see that distinct ergodic measures must have disjoint support. If μ and ν are ergodic measures and A set with $\mu(A) \neq \nu(A)$, then by (iii) we have

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T P_t I_A(x) dt = \mu(A) \quad \mu - \text{a.s.}$$

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T P_t I_A(x) dt = \nu(A) \quad \nu - \text{a.s.},$$

which implies μ and ν are singular. Ergodic behaviour is guaranteed, if one proves uniqueness of invariant measure.

Theorem 1.1.3. *If semigroup P_t has unique invariant measure μ, then this measure is ergodic.*

Proof. Assume μ is not ergodic, i. e. there is set B with $\mu(B) \in (0, 1)$ satisfying

$$P_t I_B = I_B \ \mu\text{-a.s. } \forall t > 0.$$

Define measure $\hat{\mu}(A) = \frac{1}{\mu(B)}\mu(A \cap B)$ for set A. We prove this measure is also invariant for P_t, which is contradiction to uniqueness assumption. We compute

$$P_t^*\hat{\mu}(A) = \int_{\mathcal{S}} P_t(x, A) d\mu(x) = \frac{1}{\mu(B)} \int_B P_t(x, A) d\mu(x) =$$

$$\frac{1}{\mu(B)} \int_B P_t(x, A \cap B) d\mu(x) + \frac{1}{\mu(B)} \int_B P_t(x, A \cap B^C) d\mu(x).$$

Since $P_t I_B = I_B$ μ-a.s., we have

$$P_t(x, A \cap B^C) \leq P_t(x, B^C) = 0 \ \mu\text{-a.s. for } x \in B$$

$$P_t(x, A \cap B) \leq P_t(x, B) = 0 \ \mu\text{-a.s. for } x \in B^C.$$

By invariance of μ it thus follows

$$P_t^*\hat{\mu}(A) = \frac{1}{\mu(B)} \int_B P_t(x, A \cap B) d\mu(x) = \frac{1}{\mu(B)} \int_{\mathcal{S}} P_t(x, A \cap B) d\mu(x) =$$

$$\frac{1}{\mu(B)} \int_{\mathcal{S}} I_{A\cap B}(x) d\mu(x) = \frac{1}{\mu(B)} \mu(A \cap B) = \hat{\mu}(A),$$

which proves the theorem. □

Another concept related to uniqueness of invariant measure is its irreducibility. Markov semigroup P_t is irreducible at time t_0, if for any nonempty open set G and all $x \in \mathcal{S}$ it holds $P_{t_0}(x, G) > 0$. Markov semigroup is regular at time t_0, if all transition probabilities $P_{t_0}(x, \cdot), x \in \mathcal{S}$ are mutually equivalent. Immediately by Chapman-Kolmogorov equality we have that irreducibility or regularity at time t_0 implies this property for all time $s > t_0$. It is easy to see that irreducibility and strong Feller property implies regularity[22] pp. 42.

Observation 1.1.4. *If a Markov semigroup $P_t, t \geq 0$, is strongly Feller at time $t_0 > 0$ and irreducible at time $s_0 > 0$, then it is regular at time $t_0 + s_0$.*

The importance of irreducibility for long time behaviour of Markov systems stems from the following theorem, that has its roots in Doob work[24].

Theorem 1.1.5. *Let P_t be a stochastically continuous Markovian semigroup and μ and invariant measure with respect to P_t. If P_t is t_0 regular for some*

$t_0 > 0$, then μ is unique invariant probability measure and for arbitrary $x \in \mathcal{S}$ and $A \in \mathcal{B}(\mathcal{S})$

$$\lim_{t\to\infty} P_t(x, A) = \mu(A). \tag{1.4}$$

Proof. See[22] Theorem 4.2.1. □

In fact weak convergence result (1.4) can be strengthen to convergence in total variation norm[74]. One of the ways to prove strong Feller property for semigroup on n-dimensional Euclidean spaces is to show the existence of continuous density of the semigroup with respect to standard Lebesgue measure.

Theorem 1.1.6. *Let P_t be Markov semigroup on $\mathbb{R}^n$ with continuous density, i. e. $P_t(x, dy) = p_t(x, y)$ and $p_t(\cdot, \cdot) \in C(\mathbb{R}^n \times \mathbb{R}^n)$. Then $P_t, t \geq 0$ is strong Feller semigroup.*

Proof. Let $x_n \to x$ be given sequence in $\mathbb{R}^n$. Since

$$\int_{\mathbb{R}^n} P_t(x_n, dy) = \int_{\mathbb{R}^n} p_t(x_n, y)dy = 1$$

we can choose compact set $K \in \mathbb{R}^n$ such that for all $x_n, x \in \mathbb{R}^n$

$$\int_K^C p_t(x., y)dy < \epsilon.$$

Let $f \in \mathcal{B}_b(\mathbb{R}^n), \|f\| < \infty$ and we calculate

$$\lim_{n\to\infty} P_t f(x_n) - P_t f(x) = \int_{\mathbb{R}^n} f(y)[p_t(x_n, y) - p_t(x, y)]dy =$$

$$\lim_{n\to\infty} \int_K f(y)[p_t(x_n, y) - p_t(x, y)]dy + \int_{K^C} f(y)[p_t(x_n, y) - p_t(x, y)]dy. \tag{1.5}$$

When $y \in K$, then using $x_n \to x$ and fact that continuous function on compact set is uniformly continuous, we have estimate

$$|f(y)[p_t(x_n, y) - p_t(x, y)]| \leq \|f\|2|p_t(x, y)|.$$

Lebesgue dominated convergence theorem now implies that first integral in (1.5) is zero. The second integral we can make arbitrary small choosing suitable compact set K. Hence $\lim_{n\to\infty} P_t f(x_n) - P_t f(x) = 0$ and P_t is indeed Feller. □

The theory built so far gives us satisfying conditions to ensure the uniqueness of invariant measure. Workable condition to ensure the existence of invariant measure and possibly some explicit convergence rates towards invariant measure (1.4) gives use of Lyapunov function. To this purpose, we define the generator of Markov process.

Definition 1.1.4 (Infinitesimal generator of Markov process)**.** Let P_t be Markov semigroup on $\mathcal{S}$. We define its infinitesimal generator L as

$$Lf(x) = \lim_{t\to 0_+} \frac{P_t f(x) - f(x)}{t} \quad x \in \mathcal{S}.$$

Its domain $\mathcal{D}_L$ is the set of functions $f : \mathcal{S} \to A$ such that the limit exists for all$x \in \mathcal{S}$.

There is in fact whole theory of Hille-Yosida, how to construct Markov processes given their generator[72] vol. I,[26]. However this approach is not very suitable when dealing with diffusions, where weak solutions of SDEs (which are equivalent to martingale problem approach) gives the optimal results[77].

There is now well-established theory of Meyn and Tweedie about exponential convergence towards invariant measure for Markov processes. The discrete case results obtained in monograph[56] were translated to continuous case in[57],[58]. Beautiful argument by Hairer and Mattingly in[33] offers very short proof. Having in mind application towards diffusions, we state the result in the following way (for the precise reference see[55] Theorem 2.5 or lecture notes by Rey-Bellet[71])

Theorem 1.1.7 (Harris - Meyn - Tweedie). *Let X_t be a Markov process on $\mathbb{R}^n$ with transition probability P_t and generator L. Suppose that following hypotheses are satisfied*

(1) The Markov process is irreducible

(2) For any $t > 0$ the Markov semigroup $P_t(x, dy)$ has a density $p_t(x, y)dy$ which is a continuous function of (x, y).

Assume there exists Lyapunov function

$$V : \mathbb{R}^n - > [1, \infty), \; V(x) \xrightarrow{\|x\| \to +\infty} +\infty$$

and constants $C, c > 0$ such that

$$LV + cV \leq C. \tag{1.6}$$

Then there exists unique invariant measure μ for the process X_t and there exist constants $K, \alpha > 0$ such that ($P_t f(a) = E^a f(X_t)$)

$$\sup_{\{f : |f(x)| \leq V(x)\}} |E^a f(X_t) - \mu(f)| \leq K V(a) e^{-\alpha t}$$

for any $a \in \mathbb{R}^n$.

In the statement of the theorem one implicitly assumes that $V \in \mathcal{D}_L$.

1.2 Convergence results for martingales

Martingales are processes, whose study originated in gambling. They capture rigorously the concept of processes, where knowing the full past doesn't influence future mean value. Fortunately their study does not bring such technicalities as that of Markov processes to begin with. Our exposition here is based on lecture notes[76], or in fact any good book on stochastic processes will do.

Definition 1.2.1 (martingale)**.** A stochastic process $X = (X_t)_{t\geq 0}$ defined on stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is super(sub)martingale with respect to filtration $\{\mathcal{F}_t\}$, if $X_t \in L^1$, $\forall t \geq 0$, and

$$E[X_t|\mathcal{F}_s] \leq (\geq) X_s \text{ a.s., if } s < t.$$

If a process is both supermartingale and submartingale, then it is a martingale.

Theorem 1.2.1 (Doob martingale convergence theorem)**.** *Let $\{X_t; t \geq 0\}$ be $\mathcal{F}_t$-supermartingale (or submartingale) that is right continuous with left limits. Assume that*

$$\sup_{t\geq 0} E|X_t| < \infty. \tag{1.7}$$

Then

$$X_\infty = \lim_{t\to\infty} X_t \textit{ exists a.s.,} \quad X \in L^1.$$

The theorem can be proved as a nice Corollary to Doob's upcrossing lemma A.2.1. Let us note that the requirement of existence of càdlàg version of super(sub)martingale isn't restrictive, as existence of this modification can be proved, see[70] pp. 63.

Proof. We do the proof for supermartingale, as submartingale case is proved by switching to $-X_t$. As we assume (1.7), it suffices to show a.s. existence of limit X_t and the integrability of limit X_∞ follows immediately by Fatou's lemma. We refer to Theorem A.2.1 for the definition of upcrossing an interval there. For given $M > 0$ we denote $I_M = \mathbb{Q}^+ \cap [0, M]$ the index set. For fixed $a < b, a, b \in \mathbb{R}$ we denote by $N([a, b], I_M, X)$ the number of upcrossing of interval $[a, b]$ by discrete martingale $X = (X_i)_{i\in I_M}$. Since

$$N([a, b], I_M, X) = \sup_{\{J_M \subset I_M; |J_M| < +\infty\}} N([a, b], J_M, X),$$

ineaulity in Doob's upcrossing lemma (A.1) implies that

$$(b-a)E\left[N([a,b],I_M,X)\right] \leq a + \sup_{t\geq 0} E[|X_t|] < \infty.$$

By passing $M \to \infty$ we deduce

$$N([a,b],Q^+,X) < \infty \text{ a.s.}$$

Thence the set

$$\Omega_0 = \cap_{\{a<b;a,b\in\mathbb{Q}\}}\{N([a,b],Q^+,X) < \infty\}$$

has measure one. For $\omega \in \Omega_0$ $X_q(\omega)$ converges, otherwise there would have to be infinite number of crossing of some interval $[a,b]$,$a,b \in \mathbb{Q}^+$ where

$$\liminf X_q(\omega) < a < b < \limsup X_q(\omega).$$

Thus $X_q \to X_\infty$ and this in turn implies the convergence of X_t, because for given $\epsilon > 0$ one finds $q_0 \in Q$ such that

$$|X_q - X_\infty| < \frac{\epsilon}{2}, \ \forall q \geq q_0.$$

By the right contunuity of process X_t we conclude that

$$|X_t - X_\infty| < \epsilon, \ \forall t \geq q_0.$$

□

2

Interacting diffusions - finite lattice case

2.1 Introduction

The study of interacting particle systems has a long and profound history, as is well evidenced by excellent monographs[53] or[40]. Initially motivated by the problems of statistical physics, the field has grown into an important area of Markov processes in itself with interesting problems and rich interplay with other subjects.
We investigate continuous spin systems with a diffusion particle on each site. Most results establishing ergodicity properties for interacting particle systems with unbounded state space are tied with the use of functional inequalities, see[30]. As for the diffusions, there has been two independent successful approaches to this problem in the 1990s, one by Zegarliński[82] and other by Da Prato and Zabczyk[21], each to their merit and deficiencies. The approach in[82] constructs the desired semigroup using finite dimensional approximations and ergodicity results are established via log Sobolev inequality, while more probabilistic approach in[21] uses the theory of SDEs on Hilbert spaces

for construction and ergodicity is tied with the dissipativity properties of underlying operators.
Both these works essentially cover only elliptic case. The question how to address some subelliptic situation has been resolved under suitable condition in[25] using analytic techniques (very recently the results were extended to cover even broader class of operators in[41] and[42]). Because in such cases even ergodicity of the finite system is highly non-trivial, important part of the result lies in conquering this problem.
Here we present a new probabilistic approach to investigate these issues. The results obtained go in some way successfully beyond Hilbert space methods of[21]. We can cover degenerate multiplicative noise as we show in the case of Heisenberg group (or Grushin plane). However we cannot prove the uniqueness of invariant measure, let alone convergence towards it. Notice however that such results usually require some assumptions about degeneration of the interactions, they should be tied with the condition on weights of the space the system live in, see[21] for example. Therefore it appears even probable, that under assumptions we work, the uniqueness of invariant measure for the system does not hold.

2.1.1 Outline of the strategy

The setting is the following; assume we have a space $(\mathbb{R}^n)^{\mathbb{Z}^d}$, the dynamics of the system can then formally be described by the operator of the form

$$\sum_{i\in\mathbb{Z}^d} \mathcal{A}_i + q_i\mathcal{B}_i, \tag{2.1}$$

where each $\mathcal{A}_i$ is the second order differential operator acting on $\mathbb{R}^n$ and on i-th coordinate of the lattice $\mathbb{Z}^d$, and $\mathcal{B}_i$ is the first order operator acting on i-th coordinate. We assume that we have interactions q_i only in drift term and they are of finite range.
We construct the infinite dimensional process using finite dimensional approximations by solving appropriate stochastic differential equations. Of course such approach is well known and nothing new in the field, see e.g.[35],[28].

The main novelty of our approach in comparison with these mentioned lies in two facts - we use martingale problem as a concept of solution, which allows us to bypass strong boundedness of coefficients assumption in[35], secondly we benefit from now well established Meyn-Tweedie 1.1.7 theory to prove ergodicity results in finite dimension.
In this chapter we give a proof of finite dimensional results, that is we show that for diffusions on finite product space one has exponential convergence towards invariant measure. Chapter three is devoted to results about infinite lattice. Using tightness arguments we construct the process corresponding to (2.1) as a solution to martingale problem. The key and most technical part follows, where we show under additional assumptions about interaction functions that the limit of our approximations is unique. This result is utilized to prove that our constructed process is genuine Markov process. The existence of the invariant measure for the constructed process is established in the end.
For clarity and brevity of exposition we illustrate our techniques with the specific example of the operators corresponding to Heisenberg group. However it should be noted, that many parts of our results are independent of the specific diffusions considered, so in the last section we also mention some other natural situation that can be dealt using our methods.
The model studied in Section 2 is one of fundamental examples of semigroup with subelliptic generator (associated to the simplest nontrivial nilpotent Lie algebra) where one has hypoellipticity in finite dimensions and a lot of other informations (as e.g. heat kernel estimates, short and long time behaviour), but not much was known in infite dimensions. Also, one should note, that it is natural generalization of Ornstein-Uhlenbeck process into Heisenberg group setting.

2.1.2 Statement of the results and strategy of the proof

Let $\mathbb{H} \cong \mathbb{R}^3 = (x, y, z)$ be the Heisenberg group (for the detailed treatment of Heisenberg group as an example of Stratified Lie group see[14], for nice and brief account of the relation to the matrix Heisenberg group see[4]) and X, Y

the generators of Lie algebra on $\mathbb{H}$, i. e.

$$X = \partial_x - \frac{1}{2} y \partial_z$$

$$Y = \partial_y + \frac{1}{2} x \partial_z.$$

We denote $D = x\partial_x + y\partial_y + 2z\partial_z$ (so that $[X, D] = X$, $[Y, D] = Y$) the so-called dilation operator.
We consider the product space $(\mathbb{R}^3)^{\mathbb{Z}^d}$, i. e. so-called continuous spin system where we have a copy of Heisenberg group at every point. We study the existence and long time behaviour of diffusion associated formally with the operator

$$L = \sum_{i \in \mathbb{Z}^d} \mathcal{L}_{\lambda_i} + q_{x_i} X_i + q_{y_i} Y_i, \tag{2.2}$$

where $X_{\cdot_i}$ is the vector field acting on the i-th coordinate, $q_{\cdot_i}$ is the interaction function with finite range, i. e. function whose value depends on all points $i \in \mathbb{Z}^d$ within some fixed length $0 < r < \infty$ counted with maximum metric. $\mathcal{L}_{\lambda_i} = X_i^2 + Y_i^2 - \lambda_i D_i$ and λ_i are positive constants.
Very important throughout our work is the notion of cylindrical function.

Definition 2.1.1 (Cylindrical function)**.** Let $f : X \to \mathbb{R}$, where $X \subset \mathbb{R}^\infty, \dim X = \infty$. We say f is cylindrical function, provided $\exists u \in \mathbb{N}$ s. t. $\exists g : \mathbb{R}^u \to \mathbb{R}$, so that

$$f(x_1, \ldots, x_u, \ldots) = g(x_1, \ldots, x_u).$$

If g has compact support, we say f is compactly supported cylindrical function.

While the operator (2.2) is formal, its action on cylindrical function is defined in a rigorous way. Denote $\Phi_f \subset \mathbb{Z}^d$ the corresponding subset for cylindrical function f, i. e. $f(\cdot) = g(\cdot)$, $g : (\mathbb{R}^3)^{\Phi_f} \to \mathbb{R}$.Then

$$Lf = L^f g = \sum_{i \in (\mathbb{R}^3)^{\Phi_f}} (\mathcal{L}_{\lambda_i} + q_{x_i} X_i + q_{y_i} Y_i) g.$$

The interaction function $q_{\cdot_i}$ is more precisely following function

$$q : (\mathbb{R}^3)^{\Pi_i} \to \mathbb{R},\ \text{where } \Pi_i = \{j \in \mathbb{Z}^d :\ \|j - i\|_{\max} \le r\}, |\Pi_i| = (2r+1)^d.$$

Π_i is therefore the set of all points that i-th particle interacts with. It is useful to distinguish another subsets of $\mathbb{Z}^d$, the interaction boxes

$$\Xi_n = \{i \in \mathbb{Z}^d :\ \max_{j \le d} |i_j| \le nr\}, |\Xi_n| = (2nr+1)^d.$$

We introduce metric space $S = (S, \rho), S \subset (\mathbb{R}^3)^{\mathbb{Z}^d}$, which is given by

$$S = \{a \in (\mathbb{R}^3)^{\mathbb{Z}^d} :\ \sum_{i \in \mathbb{Z}^d} \|a_i\|_{\mathbb{H}}^8 w(i) < +\infty\},$$

where $w(i) > 0$ are positive weights. For point $a \in S$ we set $\|a\|_S = \sqrt[8]{\sum_{i \in \mathbb{Z}^d} \|a_i\|_{\mathbb{H}}^8 w(i)}$, so that

$$\rho(a, b) = \sqrt[8]{\sum_{i \in \mathbb{Z}^d} \|a_i - b_i\|_{\mathbb{H}}^8 w(i)}$$

.

Assumptions. To achieve our results it turns out that we need to impose the following six hypotheses :

- **(H1)** $\exists C > 0 :\ \sup_{u \in (\mathbb{R}^3)^{(2r+1)^d}} |q_{\cdot_i}(u)| \le C,\ i \in \mathbb{Z}^d$
- **(H2)** $\sup_{u \in (\mathbb{R}^3)^{(2r+1)^d}} \sum_{j=1}^{(2r+1)^d} |\frac{\partial q_{\cdot_i}}{\partial_j}(u) u_{\cdot_i}| + |\frac{\partial q_{\cdot_i}}{\partial_j}(u)| \le C,\ i \in \mathbb{Z}^d$
- **(H3)** $\inf_{i \in \mathbb{Z}^d} \lambda_i > 0, \sup_{i \in \mathbb{Z}^d} \lambda_i < \infty.$
- **(H4)**$\sum_{i \in \mathbb{Z}^d} w(i) < +\infty,\ w(i) > 0, i \in \mathbb{Z}^d$
- **(H5)** $\exists v(i) > 0, i \in \mathbb{Z}^d, \sum_i v(i) < +\infty, \sum_i \frac{w(i)}{v(i)} < +\infty$
- **(H6)** $\exists \delta \in (0,1)\ \exists K > 0$ s. t.

$$w(j) \ge \frac{K}{i!^{1-\delta}}\quad j \in \Xi_i \setminus \Xi_{i-1},\ i \in \mathbb{N}.$$

Remark. The first assumption tells us that we work only with bounded interactions, which is often the limitation, see[21],[35].
Second hypothesis may seem bit artificial and is dictated to us by the method we use in the proofs. The second condition means that the rate of change of interaction is bounded. The boundedness of the first term we could interpret in a way, that if the value of particle is large, then the interaction strength does no longer change its value. The third hypothesis is natural, since even in the case of single diffusion, to get the existence of invariant measure λ must be positive. So this just ensures that the size of the interaction doesn't degenerate, nor does it grow.
(H4) and **(H5)** ensure that the space is large enough to accommodate some non-trivial process and invariant measure. It is well known, that some conditions of this form are necessary, the exact form of **(H5)** is interesting and comes as a result of our methods.
On the other hand, **(H6)** limits the size of initial configurations for which we can prove some long time behaviour. Notice that restrictions of our method are mild as we allow even factorial growth, since in[21] pp. 10 the authors must assume at most polynomial growth.

We can now summarize our results as follows. While this operator is just formal, from the definition it is clear, how its action of smooth cylindrical function looks like.

Theorem 2.1.1 (Inifinite dimensional results)**.** *Let L be the operator given by (2.2). Assume that **(H1) - (H6)** hold. Then for any $a \in S$ there exists probability measure P^a on $\Omega = C([0,\infty), S)$ such that for the canonical process $A_t(\omega) = \omega_t$ on Ω we have $P^a(A_0 = a) = 1$ and the process*

$$f(A_t) - f(A_0) - \int_0^t Lf(A_u)du$$

is martingale for $f \in C_c^{2,Cyl}$ under the measure P^a. The pair (A_t, P^a) is a Markov process and there exist an invariant measure ν for the semigroup

$$P_t f(a) = E^a f(A(t)), a \in S.$$

The Theorem consists of several non-trivial ingredients. The existence of solution to the martingale problem is proved in Theorem 3.1.7, Markov property in Theorem 3.2.6 and the existence of invariant measure for the model is proved in Chapter 3.3.
To reach these results, we firstly proceed by investigating the case of diffusion on Heisenberg group. Concretely we analyse the asymptotic behaviour of the Markov process on $\mathbb{R}^3$ associated with the operator

$$\mathcal{L} = X^2 + Y^2 - \lambda D + q_x X + q_y Y.$$

Under suitable assumptions on q's the process can be constructed by ordinary Itō stochastic equation and using the theory of Meyn and Tweedie (1.1.7) we establish exponential convergence in the total variation norm to the invariant measure in section 2. This result can be immediately translated to the exponential ergodicity of diffusion on $(\mathbb{R}^3)^u$, $u \in \mathbb{N}$ with the generator

$$\sum_{i=1}^{u} \mathcal{L}_{\lambda_i} + q_{x_i} X_i + q_{y_i} Y_i.$$

We prove explicitly the following result.

Theorem 2.1.2 (Finite Dimensional results)**.** *Let $(\mathbb{R}^3)^u$, $u \in \mathbb{N}$ be the state space and consider the operator*

$$L_u = \sum_{i=1}^{u} \mathcal{L}_{\lambda_i} + q_{x_i} X_i + q_{y_i} Y_i, \tag{2.3}$$

*under the corresponding assumptions **(H1), (H3)**. Let us denote A^u the diffusion corresponding to the operator (2.3), i. e. the unique solution to the Itō SDE with coefficients*

$$\begin{aligned} b =& (q_{1,x} - \lambda_1 x_1, q_{1,y} - \lambda_1 y_1, -2\lambda_1 z_1 + \frac{1}{2}(q_{1,y} x_1 - q_{1,x} y_1), \dots \\ & \dots, q_{u,x} - \lambda_u x_u, q_{u,y} - \lambda_u y_u, -2\lambda_u z_u + \frac{1}{2}(q_{u,y} x_u - q_{u,y} y_u)) \end{aligned}$$

$$\sigma = \begin{pmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & M_u \end{pmatrix}, \ \textit{where } M_i = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \\ \frac{-y_i}{\sqrt{2}} & \frac{x_i}{\sqrt{2}} \end{pmatrix}.$$

There exists unique invariant measure μ_u for the process A^u. For the function $W_u^k = 1 + \sum_{i=1}^u v(i)((x_i^2 + y_i^2)^2 + z_i^2)^k$, $k \in \mathbb{N}$, where $v(i) > 0, \sum_i v(i) < \infty$, there exist constants $c_k > 0$ and $C_k > 0$ such that

$$L_u W_u^k(a_u) \leq -c_k W_u^k(a_u) + C_k \ \forall a_u \in (\mathbb{R}^3)^u. \tag{2.4}$$

In addition there exist constant $K_u^k, \alpha_u^k > 0$, such that the following

$$\sup_{\{f \in \mathcal{B}(\mathbb{R}^{3u}) \, : \, \|f\|_\infty \leq 1\}} |E^a f(A^u(t)) - \mu_u(f)| \leq K_u^k W_u^k(a) e^{-\alpha_u^k t} \tag{2.5}$$

holds for any $a \in (\mathbb{R}^3)^u$.

Next we consider an exhausting sequence $\Lambda_n \subset\subset \mathbb{Z}^d, \Lambda_n \nearrow \mathbb{Z}^d$ of the lattice. On each $(\mathbb{R}^3)^{\Lambda_n}$ we construct diffusion A^n that its generator extends the operator

$$L_n = \sum_{i \in \Lambda_n} \mathcal{L}_{\lambda_i} + q_{x_i}^n X_i + q_{y_i}^n Y_i.$$

Unfortunately unlike in [35] we are in a situation with unbounded coefficients, so we are unable to prove the limit of approximations in the strong sense. Nevertheless we show tightness in appropriate weighted space S, $S \subset (\mathbb{R}^3)^{\mathbb{Z}^d}$, i. e. we are able to prove that the distributions of the processes $\tilde{A}^n = (A^n, 0_{i \in \mathbb{Z} \setminus \Lambda_n})$ form a tight sequence in $\Omega = C([0, \infty), S)$. From tightness follows the construction of family of measures $\{P^a, a \in S\}$ such that canonical process on Ω solves the martingale problem for (2.2). Our results are not completely satisfactory since we do not address the uniqueness of martingale problem for the operator (2.2).
Nevertheless under additional assumptions we can prove that our approximation procedure yields a unique measure. This is used to show that canonical process is a proper Markov process under constructed measure. Furthermore exploiting the results obtained for bounded lattice we prove the existence of

invariant measure for the unbounded lattice.
In certain aspects therefore - such as requiring no further assumptions on λ in relevant examples - our results compare favourably to the ones in [25], [41]. However it should be noted that our methods are only able to handle bounded interactions q's and we also work with much simpler generators than the authors in the above mentioned articles. One could also argue that our proofs are bit simpler, although that perhaps depends more on the background of the reader as they are still very technical.

2.2 Finite dimensional case

In this section the case of diffusion on $\mathbb{R}^3$, respectively the diffusion on finite product space $(\mathbb{R}^3)^u$ is investigated. With our assumptions the construction is immediate and we use Meyn-Tweedie theory to study its long time behaviour.
We start by analyzing the diffusion on $\mathbb{R}^3$ associated with the second order operator

$$\mathcal{L} = X^2 + Y^2 - \lambda D + q_x X + q_y Y. \tag{2.6}$$

We will work under the following assumptions **(B1)** :

- $q_x, q_y \in C^\infty(\mathbb{R}^3, \mathbb{R}), \lambda > 0$
- $\exists\ C > 0\ :\ ||q_x||_\infty \vee ||q_y||_\infty \leq C$

Under these assumptions we can construct the diffusion as a solution to the SDE

$$dA(t) = b(A_t)dt + \sigma(A_t)dW_t.$$

Elementary computations with vector fields and matrices reveal that the coefficients can be chosen as

$$
\begin{gathered}
b = \left(q_x - \lambda x, q_y - \lambda y, -2\lambda z + \frac{1}{2}(q_y x - q_x y)\right) \\
\sigma = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \\ \frac{-y}{\sqrt{2}} & \frac{x}{\sqrt{2}} \end{pmatrix} \longrightarrow \frac{1}{2}a = \frac{1}{2}\sigma\sigma^* = \begin{pmatrix} 1 & 0 & \frac{-y}{2} \\ 0 & 1 & \frac{x}{2} \\ \frac{-y}{2} & \frac{x}{2} & \frac{1}{4}(x^2+y^2) \end{pmatrix}.
\end{gathered} \tag{2.7}
$$

We aim to show, that A_t satisfies the condition of Theorem 1.1.7. Every verification of this result is non-trivial and depends on deep results about diffusions in $\mathbb{R}^u$. In the remainder of the section we show that the process A given by SDE with the coefficients (2.7) indeed satisfies the condition of the above theorem. The existence and smoothness of transition probability density is the immediate consequence of the Hörmander theorem in probabilistic settings. The version that is sufficient for our purposes was first established following Hörmander work[36] in[37].

Theorem 2.2.1 (Hörmander probabilistic setting, Ichihara - Kunita)**.** *Assume X_t is the unique strong solution to the Stratonovich SDE*

$$
dX_t = b(X_t)dt + \sum_{i=1}^{d} \sigma(X_t) \circ dW_t,
$$

where $b, \sigma_i, 1 \leq i \leq d \ \in C^\infty(\mathbb{R}^u, \mathbb{R})$. Suppose that the following (Hörmander) condition is satisfied

$$
dim\ (Lie\{\sigma_1, \dots, \sigma_d\}) = u \ \forall x \in \mathbb{R}^u. \tag{2.8}
$$

Then there exists probability density function $P_t(x, dy) = p_t(x, y)dy$ such that $p_t(x, y) \in C^\infty((0, \infty), \mathbb{R}^u, \mathbb{R}^u)$.

Remark. The condition (2.8) is a specific instance of the widely known parabolic Hörmander condition, which ensures the hypoellipticity of $\partial_t - \mathcal{L}^*$ ($\mathcal{L}^*$ being the formal L^2-adjoint of $\mathcal{L}$), which in turn implies existence of a smooth density.

In our case (2.7) the drift in the Stratonovich form is actually the same as in Itō form. In any case the Lie algebra generated by the diffusion satisfies the Hörmander condition as elementary computation reveals that $[X, Y] = \partial_z$ and consequently

$$\dim\left(\text{Lie}\left\{\left(\sqrt{2}, 0, \frac{-y}{\sqrt{2}}\right), \left(0, \sqrt{2}, \frac{x}{\sqrt{2}}\right)\right\}\right) = 3.$$

Thence according to the above cited theorem we have the smoothness of transition probability density for (2.7).

To investigate the irreducibility of diffusion, we would like to use Stroock - Varadhan support theorem[78], so the question is whether that we can solve the corresponding control problem. The version that accounts for unbounded coefficients we use, was proved in[31].

Let F be the subset of the absolutely continuous functions $v : [0, t] \to \mathbb{R}^d$ with $v(0) = 0$ such that F contains every infinitely differentiable function from $[0, t]$ to $\mathbb{R}^d$ vanishing at zero. For the ordinary differential equation

$$\begin{aligned} \dot{x}^v(t) &= b(x^v(t)) + \sum_{i=1}^{d} \dot{v}_i(t)\sigma_i(x^v(t)) \\ x^v(0) &= x_0 \ \in \mathbb{R}^u \end{aligned} \tag{2.9}$$

we denote $\mathcal{O}(t, x_0) = \{y \in \mathbb{R}^u : x^v(t) = y, v \in F\}$ its orbit.

Theorem 2.2.2 (Stroock - Varadhan support theorem,[31])**.** *Let X_t be the solution to the Stratonovich SDE*

$$dX_t = b(X_t)dt + \sum_{i=1}^{d} \sigma_i \circ dW, \ X(0) = x, \tag{2.10}$$

where the coefficients satisfy linear growth assumptions, see (A.1.1), b is Lipschitz and $\sigma_i, 1 \leq i \leq d$ are smooth with bounded derivatives. Let P_t be the transition probability function related to (2.10) and $\mathcal{O}(t, x)$ be the orbit to the corresponding equation (2.9). Then supp $P_t(x, \cdot) = \overline{\mathcal{O}(t, x)}$.

Lemma 2.2.3. *Let P_t be the transition function for the equation (2.7). Then* *supp* $P_t(x,\cdot) = \mathbb{R}^3$ *for any* $t > 0$ *and* $x \in \mathbb{R}^3$.

Proof. We make of use the classical Girsanov transformation A.2.2 to simplify the control problem. Concretely the statement that the support of diffusions X_t, Y_t

$$dX_t = b(X)dt + \sigma(X)dW$$

$$dY_t = \tilde{b}(Y)dt + \sigma(Y)dW, \tag{2.11}$$

where σ and b are as in (2.7) and

$$\tilde{b} = (-\lambda x, -\lambda y, -2\lambda z)$$

is the same, if there is $u : \mathbb{R}^3 \to \mathbb{R}^{2\times 1}$ such that

$$\sigma u = b - \tilde{b}.$$

However it is easy, since $b - \tilde{b} = (q_x, q_y, \frac{1}{2}(q_y x - q_x y))$ and hence

$$\begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \\ \frac{-y}{\sqrt{2}} & \frac{x}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{q_x}{\sqrt{2}} \\ \frac{q_y}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} q_x \\ q_y \\ \frac{1}{2}(q_y x - q_x y) \end{pmatrix}.$$

Therefore to establish the theorem it suffices to prove the irreducibility of transition function corresponding to (2.11). Since the equation (2.11) satisfies the Theorem 2.2.2, we only need to prove that the closure of orbit for system

$$\begin{aligned} \dot{x} &= \sqrt{2}\dot{u}_1 - \lambda x \\ \dot{y} &= \sqrt{2}\dot{u}_2 - \lambda y \\ \dot{z} &= -\frac{y}{\sqrt{2}}\dot{u}_1 + \frac{x}{\sqrt{2}}\dot{u}_2 - 2\lambda z \end{aligned} \tag{2.12}$$

for $u \in H$ is full space, i. e. to show that from any starting point (x_0, y_0, z_0) we can choose such $u \in H$ that $x(t) = x_t, y(t) = y_t, z(t) = z_t$, where $(x_t, y_t, z_t) \in \mathbb{R}^3$ are prescribed ending points. If we simply choose control $\dot{u}_1(s) = as + b$, $\dot{u}_2(s) = cs + d$, then the problem (2.12) is reduced to solving three linear equations with four parameters, so the Lemma is proved. $\square$

The proof of existence of Lyapunov function for the operator (2.6) satisfying (1.6) is elementary, albeit bit tedious.

Lemma 2.2.4. *Let $\mathcal{L}$ be the operator defined by (2.6) under the assumptions* ***(B1)****. For the function $V^k = ((x^2+y^2)^2+z^2)^k$, $k \in \mathbb{N}$, there exist constants $c_k, C_k > 0$ such that*

$$\mathcal{L}V^k + c_k V^k \leq C_k \ \forall (x,y,z) \in \mathbb{R}^3. \tag{2.13}$$

Proof. We first compute the case for V^k, $k=1$ (and omit the index in such case) and then proceed to general k. Using that $V_{xz} = V_{yz} = 0$ we calculate

$$\begin{aligned}
&\mathcal{L}V + cV = V_x(q_x - \lambda x) + V_y(q_y - \lambda y) + V_z(-2\lambda z + \frac{1}{2}(q_y x - q_x y)) \\
&+ V_{xx} + V_{yy} + V_{zz}\frac{1}{4}(x^2+y^2) = 4x(x^2+y^2)(q_x - \lambda x) \\
&+ 4y(x^2+y^2)(q_y - \lambda y) + 2z(-2\lambda z + \frac{1}{2}(q_y x - q_x y)) + 16(x^2+y^2) \\
&+ \frac{1}{2}(x^2+y^2) + cx^4 + cy^4 + 2x^2y^2c + cz^2 \\
&\leq (x^4+y^4+z^2+2x^2y^2)(c - 4\lambda) + o(x^4) + o(y^4) + o(z^2).
\end{aligned} \tag{2.14}$$

To obtain last inequality one uses bounds for $q_{\cdot}$'s and then Young inequality, e. g.

$$|Czx| \lesssim |z|^{\frac{3}{2}} + |x|^3.$$

The resulted inequality obviously implies that for any $\lambda > 0$ we can choose $c > 0$ in such way, that $\mathcal{L}V + cV$ is bounded. For V^k we get

$$\begin{aligned}
&\mathcal{L}V^k + c_k V^k = V^{k-2}k\big(VV_x(q_x - \lambda x) + VV_y(q_y - \lambda y) \\
&+ VV_z(-2\lambda z + \frac{1}{2}(q_y x - q_x y)) + (k-1)V_x^2 + VV_{xx} \\
&+ (k-1)V_y^2 + VV_{yy} + \frac{1}{4}(x^2+y^2)(VV_{zz} + (k-1)V_z^2) \\
&- y(k-1)V_xV_z + x(k-1)V_yV_z + \frac{c_k}{k}V^2\big).
\end{aligned} \tag{2.15}$$

In a similar manner as we obtained (2.14), (2.15) can be estimated as

$$\mathcal{L}V^k + c_kV^k \le V^{k-2}k\big((\frac{c_k}{k} - 4\lambda)(x^8 + y^8 + z^4)$$
$$+ x^4y^4(\frac{4c_k}{k} - 8\lambda) + o(x^8) + o(y^8) + o(z^4)\big).$$

Notice that we not only proved boundedness of $\mathcal{L}V^k + c_kV^k$, but even obtained

$$\mathcal{L}V^k + c_kV^k \xrightarrow{\|(x,y,z)\| \to +\infty} -\infty.$$

□

The Meyn - Tweedie theory as stated in Theorem 1.1.7 now ensures exponential convergence to equilibrium for diffusion corresponding to the operator (2.6).

Proof of Theorem 2.1.2. The proof does not contatin adidtional ideas compared to what we just showed in the case of diffusion on $\mathbb{R}^3$. This is thanks to the structure of our coefficients b and σ and assumptions **(H1), (H3)**. As for the support problem, the situation is pretty much the same as in Lemma 2.2.3, and the smoothness of transition probability follows again immediately from Hörmander type theorem 2.2.1.
It remains to show that for W^k there exist constants $c_k, C_k > 0$ such that

$$L_uW_u^k + c_kW_u^k \le C_k$$

holds uniformly regardless of $u \in \mathbb{N}$. By denoting $V_i^k = ((x_i^2 + y_i^2)^2 + z_i^2)^k$ and $\mathcal{L}_i = X_i^2 + Y_i^2 - \lambda_i D_i + q_{x_i}X_i + q_{y_i}Y_i$, we can write

$$L_uW_u^k + c_kW_u^k = c_k + \sum_{i=1}^{u} w(i)(\mathcal{L}_iV_i^k + c_kV_i^k). \tag{2.16}$$

The analysis of expression $\mathcal{L}_iV_i^k + c_kV_i^k$ was done in previous Lemma 2.2.4. Notice that thanks to the assumption $\inf_{i\in\mathbb{Z}^d} \lambda_i > 0$ and the fact that bound for $q_{\cdot}$'s is uniform, the c_k can be chosen in such way, that the following is true

$$\exists \tilde{C}_k > 0 \ : \forall_{1\leq i\leq u} \ \ \mathcal{L}_i V_i^k + c_k V_i^k \leq \tilde{C}_k.$$

We insert this into (2.16) and use hypothesis $\sum_{i=1}^{\infty} w(i) < \infty$ to infer the desired bounds

$$\begin{aligned} L_u W^k + c_k W^k &\leq c_k + \sum_{i=1}^{u} u(i)\tilde{C}_k \\ &\leq c_k + \tilde{C}_k \sum_{i=1}^{\infty} w(i) := C_k < +\infty. \end{aligned}$$

Hence indeed the Theorem 1.1.7 can be applied to prove the statement. $\square$

Remark. It should be noted that the constants in the formula (2.5) cannot be chosen in such a way, that they would be independent of the dimension, even though the constants in (2.4) are. It cannot be expected that one could prove the convergence in the total variation norm in the infinite dimension for interacting particle system, at least for interesting systems. Let us add a simple reason for this fact.

Observation 2.2.5. *Let ϖ and ϱ be two probability measures on $\mathbb{R}$, such that $\varpi \neq \varrho$. Then for the product measures ϖ^u and ϱ^u on $\mathbb{R}^u$ it holds that*

$$\|\varpi^u - \varrho^u\|_{TV} \xrightarrow{u\to\infty} 1,$$

where $\|\cdot\|_{TV}$ means the total variation norm.

Proof. As $\varpi \neq \varrho$, there exists $f \in C_b(\mathbb{R})$, such that $\varpi f \neq \varrho f$.
Put $\epsilon = |\varpi f - \varrho f| > 0$. Define sets $A_u = \{x \in \mathbb{R}^u : |\frac{1}{u}\sum_{i=1}^{u} f(x_i) - \varpi f| < \frac{\epsilon}{2}\}$.
Weak Law of large numbers asserts $\varpi^u(A_u) \to 1$, while $\varrho^u(A_u) \to 0$. $\square$

Therefore even for the system without any interactions, one cannot have the constants independent of the dimension, unless the convergence to the invariant measure happens in finite amount of time.

3

Infinite system of interacting diffusions

3.1 Construction of the infinite dimensional measure

The goal of this section is to construct solution to the martingale problem associated with the operator (2.2). This is done by approximation by finite dimensional stochastic equations and using compactness arguments.

There are several papers dealing with infinite dimensional martingale problems ([2],[7],[81]) that establish uniqueness as well, but all are based in elliptic settings and none can be directly applied to our case.
The following version of Arzelà - Ascoli theorem follows easily from the general version proved in[59] Theorem 47.1.

Theorem 3.1.1 (Arzelà - Ascoli)**.** *Let Y be a complete metric space and $f_n \in C([0,\infty), Y), n \in \mathbb{N}$ sequence of equicontinuous functions. Endow $C([0,\infty), Y)$ with the topology of uniform convergence on compacts. If $\{f_n(t)\}$ is precompact in Y on a dense set of $t \in [0,\infty)$, then $\{f_n\}$ is precompact in $C([0,\infty), Y)$.*

To prove equicontinuity we use a variant of Kolmogorov continuity theorem (see[8] chap. 8 for details).

Theorem 3.1.2. *Let X^n be continuous processes taking values in some metric space (S, ρ). Suppose for any $T > 0$ there exist constants $C(T), \epsilon > 0$ and $p > 0$ such that*

$$\sup_n E\ \rho(X_s^n, X_t^n)^p \leq C(T)|t-s|^{1+\epsilon} \quad 0 \leq s \leq t \leq T.$$

Then $\{X^n\}$ is equicontinuous family of processes with probability 1.

The space on which we construct our measure is dictated to us by our Lyapunov function for (2.6), so that we will be able to utilize the uniform bound (2.5). However we also have to choose space such that the Theorem 3.1.2 will be satisfied. For the sake of completeness let us clarify, that function of V type indeed equips $\mathbb{R}^3$ with the metric.

Lemma 3.1.3. *Endow $\mathbb{R}^3$ with the following operation d :*

$$d(a,b) = \sqrt[4]{((a_x-b_x)^2+(a_y-b_y)^2)^2+(a_z-b_z)^2}.$$

$(\mathbb{R}^3, d)$ is then a metric space.

Proof. The only non-trivial part is the triangle inequality. Hence we want to prove

$$\begin{aligned} &\sqrt[4]{((a_x-b_x)^2+(a_y-b_y)^2)^2+(a_z-b_z)^2} \qquad \leq \\ &\sqrt[4]{((a_x-c_x)^2+(a_y-c_y)^2)^2+(a_z-c_z)^2} \\ &+\sqrt[4]{((c_x-b_x)^2+(c_y-b_y)^2)^2+(c_z-b_z)^2}. \end{aligned} \tag{3.1}$$

Notice that (3.1) is clearly valid if either terms on z axis are zero, or both x and y terms are zero. Therefore it remains to prove that if for $A, B, C, D, E, F \geq 0$

$$\begin{aligned} \sqrt[4]{A} &\leq \sqrt[4]{B} + \sqrt[4]{C} \\ \sqrt[4]{D} &\leq \sqrt[4]{E} + \sqrt[4]{F}, \end{aligned} \tag{3.2}$$

then

$$\sqrt[4]{A+D} \leq \sqrt[4]{B+E} + \sqrt[4]{C+F}. \tag{3.3}$$

The left side in (3.3) is clearly maximized, if the left sides in (3.2) is maximized. This happens, if we have equality in (3.2). Hence it suffices to prove

$$\sqrt[4]{(\sqrt[4]{B}+\sqrt[4]{C})^4 + (\sqrt[4]{E}+\sqrt[4]{F})^4} \leq \sqrt[4]{B+E} + \sqrt[4]{C+F},$$

but this follows from ordinary Minkowski inequality for 4 - norm on $\mathbb{R}^2$. $\square$

Remind our weighted space S

$$S = \{a \in \mathbb{H}^{\mathbb{Z}^d} : \sum_{i\in\mathbb{Z}^d} \|a_i\|_{\mathbb{H}}^8 w(i) < +\infty\}.$$

For now it suffices to assume that weights satisfy **(H4)**. From the Lemma above we can infer following usual procedure that S with the metric

$$\|a-b\|_S = \sqrt[8]{\sum_{i\in\mathbb{Z}^d} \|a_i - b_i\|_{\mathbb{H}}^8 w(i)},\ a, b \in S$$

is a complete separable metric space. Therefore $\Omega = C([0,\infty), S)$ is Polish too. Let us describe compact sets of S.

Lemma 3.1.4. *Let $M \subset S$. Assume that M is bounded and the following condition*

$$\forall \epsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall a \in M : \sum_{i=n_0}^{\infty} \|a_i\|_{\mathbb{H}}^8 w(i) < \epsilon.$$

Then M is precompact in S.

Proof. We show that from any sequence $\{a^n\}$ one can extract a Cauchy sequence. By assumptions for a given $\epsilon > 0$ we find n_0, so we control the rest of the sequence, and on the first $n_0 - 1$ coordinates we simply choose a Cauchy sequence step by step, which is possible by the boundedness assumption. $\square$

3.1.1 Moments estimates and tightness of approximations

Let $\Lambda_n, |\Lambda_n| = N < +\infty$ be the exhausting sequence of $\mathbb{Z}^d$, i. e. $\Lambda_{n+1} \supseteq \Lambda_n$, $\bigcup_n \Lambda_n = \mathbb{Z}^d$. We wish to construct martingale solution for the operator

$$L = \sum_{i \in \mathbb{Z}^d} \mathcal{L}_{\lambda_i} + q_{x_i} X_i + q_{y_i} Y_i \ . \tag{3.4}$$

Suppose we have maximum metric on $\mathbb{Z}^d$ and we assume there exists constant $r > 0$ such that $q_{\cdot_i}$ depends only on neighbours within distance r. More precisely we assume about interaction functions q's **(H1)**, the constants λ_i are subjected to **(H3)**.

On each space $(\mathbb{R}^3)^N$ we consider diffusion A^n with generator that coincides on $C_c^2((\mathbb{R}^3)^{\Lambda_n})$ functions with

$$L_n = \sum_{i \in \Lambda_n} L_{\lambda_i} + q^n_{x_i} X_i + q^n_{y_i} Y_i. \tag{3.5}$$

Remember that the interaction functions $q_{\cdot_i}$ are the functions of all neighbours in distance r (we assume the max metric). Therefore in case point $i \in \Lambda_n$ does not contain all neighbours in distance r, we must redefine $q^n_{\cdot_i}$, so that its behaviour is dependent only on "available" points, i. e. on set

$$\{j \in \mathbb{Z}^d : j \in \Pi_i = \{j \in \mathbb{Z}^d : |j - i|_{\max} \leq r\} \cap \Lambda_n\}.$$

The redefinition is arbitrary, we only have to demand interaction functions still obey **(H1)**. If

$$\Pi_i \subset \Lambda_n, \text{ we put } q^n_{\cdot_i} = q_{\cdot_i}. \tag{3.6}$$

Notice that if A^n is a solution to the corresponding stochastic equation

$$dA^n = b^n(A^n)dt + \sigma^n(A^n)dW_t \text{ on } (\mathbb{R}^3)^{\Lambda_n},$$

then whenever $\Pi_i \subset \Lambda_n$ (i. e. all points that i-th particle interacts with), the b^n_i no longer depends on n and we can just write in such cases $b^n_i = b_i$. Naturally the situation with σ is even simpler, since the coefficients are independent of the n, there are no interactions in the diffusion term. Setting

$$\tilde{A}^n = (A^n, 0_{i\in\mathbb{Z}^d\setminus\Lambda_n}),$$

each $\tilde{A}^n(t)$ has values in S and therefore $\tilde{A}^n$ lives in $\Omega = C([0,\infty), S)$.

Lemma 3.1.5. *Let $a \in S$. For $n \in \mathbb{N}$ define A^n as above with initial condition $A^n(0) = \pi_{\Lambda_n}(a)$ and subsequently define $\tilde{A}^n$. Assume **(H1), (H3), (H4)**. Let $T > 0$ be given. Then there exist constants $C(T) > 0$*

$$\sup_n \ \forall_{0\le s\le t\le T} \ E\|\tilde{A}^n(t) - \tilde{A}^n(s)\|_S^8 \le C(T)|t-s|^2 \tag{3.7}$$

$$\forall\delta > 0 \ \forall t \ge 0 \ \exists N_0(t,\delta) \ : \ \sup_n \ E \sum_{i=N_0(t)+1}^{\infty} \|A_i^n(t)\|_{\mathbb{H}}^8 u(i) < \delta. \tag{3.8}$$

Proof. First notice that the assumptions lead to the existence of constant K such that (b^n, σ^n being the coefficients of SDE for A^n)

$$\begin{aligned} &|b_{i,x}^n(a)| \vee \|\sigma_{i,x}^n(a)\|_{\mathbb{R}^{2N}} \le K(1+|a_{i,x}|) \\ &|b_{i,y}^n(a)| \vee \|\sigma_{i,y}^n(a)\|_{\mathbb{R}^{2N}} \le K(1+|a_{i,y}|) \\ &|b_{i,z}^n(a)| \vee \|\sigma_{i,z}^n(a)\|_{\mathbb{R}^{2N}} \le K(1+\sum_{j=1}^{3}|a_{i,j}|). \end{aligned} \tag{3.9}$$

Suppose $0 < s < t \le T$, we have (remind $|\Lambda_n| = N$)

$$E\|\tilde{A}^n(t) - \tilde{A}^n(s)\|_S^8 = E\sum_{i=1}^{N} \|A_i^n(t) - A_i^n(s)\|_{\mathbb{H}}^8 w(i)$$

$$= \sum_{i=1}^{N} E(((A_{i,x}^n(t) - A_{i,x}^n(s))^2 + (A_{i,y}^n(t) - A_{i,y}^n(s))^2)^2 + (A_{i,z}^n(t) - A_{i,z}^n(s))^2)^2 w(i)$$

$$\lesssim \sum_{i=1}^{N} w(i)\Big(E(A_{i,x}^n(t) - A_{i,x}^n(s))^8 + E(A_{i,y}^n(t) - A_{i,y}^n(s))^8 + E(A_{i,z}^n(t) - A_{i,z}^n(s))^4\Big). \tag{3.10}$$

The x term is now estimated using (3.9), Burkholder - Davis - Gundy and

Hölder inequalities

$$E(A^n_{i,x}(t) - A^n_{i,x}(s))^8 = E\left(\int_s^t b^n_{i,x}(A^n(u))du + \int_s^t \sigma^n_{i,x}(A^n(u))dW(u)\right)^8$$

$$\lesssim |t-s|^7 E(\int_s^t |b^n_{i,x}(A^n(u))|^8 du) + |t-s|^3 E(\int_s^t \|\sigma^n_{i,x}(A^n(u))\|^8 du)$$

$$\lesssim |t-s|^2 + |t-s| \int_s^t E|A^n_{i,x}(u)|^8 du.$$

Similarly handling the y and z we get

$$E(A^n_{i,y}(t) - A^n_{i,y}(s))^8 \lesssim |t-s|^2 + |t-s| \int_s^t E|A^n_{i,y}(u)|^8 du)$$

$$E(A^n_{i,z}(t) - A^n_{i,z}(s))^4 \lesssim |t-s|^2 + |t-s| \int_s^t \sum_{j=1}^3 E|A^n_{i,j}(u)|^4 du.$$

Individual terms we treat

$$E|A^n_{i,x}(u)|^8 = E\left||a_{i,x}| + \int_0^u b^n_{i,x}(A^n(v))dv + \int_0^u \sigma^n_{i,x}(A^n(v))dW(v)\right|^8$$

$$\lesssim |a_{i,x}|^8 + 1 + \int_0^u E|A^n_{i,x}(v)|^8 dv,$$

analogically one gets

$$E|A^n_{i,y}(u)|^8 \lesssim |a_{i,y}|^8 + 1 + \int_0^u E|A^n_{i,y}(v)|^8 dv$$

$$E|A^n_{i,z}(u)|^4 \lesssim |a_{i,z}|^4 + 1 + \int_0^u \sum_{j=1}^3 E|A^n_{i,j}(v)|^4 dv.$$

Altogether we derived existence of some constant $K(T) > 0$ such that

$$E|A^n_{i,x}(u)|^8 + E|A^n_{i,y}(u)|^8 + \sum_{j=1}^3 E|A^n_{i,j}(u)|^4 \leq K(T)(\|a_i\|^8_{\mathbb{H}} + 1)$$

$$+K(T)\int_0^u \Big(E|A^n_{i,x}(u)|^8 + E|A^n_{i,y}(u)|^8 + \sum_{j=1}^3 E|A^n_{i,j}(u)|^4\Big)du.$$

Invoking the Grönwall's inequality we can deduce existence of some constant $K_1(T) >$ such that $\forall u \in [s,t]$

$$E|A^n_{i,x}(u)|^8 + E|A^n_{i,y}(u)|^8 + \sum_{j=1}^3 E|A^n_{i,j}(u)|^4 \leq K_1(T)(1 + \|a_i\|^8_{\mathbb{H}}). \tag{3.11}$$

Hence

$$E(A^n_{i,x}(t) - A^n_{i,x}(s))^8 + E(A^n_{i,y}(t) - A^n_{i,y}(s))^8 + E(A^n_{i,z}(t) - A^n_{i,z}(s))^4$$

$$\lesssim |t-s|^2 + |t-s|^2 K_1(T)(1 + \|a_i\|_{\mathbb{H}}).$$

Inserting back to (3.10) we obtain, thanks to **(A2)** and the fact that $a \in S$ the existence of some constants $L(T), C(T) > 0$ such that

$$E\|\tilde{A}^n(t) - \tilde{A}^n(s)\|^8_S \leq \sum_{i=1}^N w(i)|t-s|^2 L(T)(1 + \|a_i\|^8_{\mathbb{H}})$$

$$\leq C(T)|t-s|^2,$$

which is what we wanted to prove in (3.7).
To prove (3.8) we simply utilize the key estimate (3.11), which gives us

$$E\sum_{N_0(t)+1}^{\infty} \|A^n_i(t)\|^8_{\mathbb{H}} w(i) \lesssim \sum_{N_0(t)+1}^{\infty} u(i)K_1(t)(1 + \|a_i\|^8_{\mathbb{H}}),$$

therefore for given $\delta > 0$ it suffices to choose $N_0(t)$ such that the sum

$$\sum_{i=N_0(t)+1}^{\infty} w(i)(1 + \|a_i\|^8_{\mathbb{H}})$$

is sufficiently small. □

Corollary 3.1.6. *Let $\tilde{A}^n$ be as in Lemma 3.1.5. Then $P \circ (\tilde{A}^n)^{-1}, n \geq 1$ is a tight sequence of measures in Ω.*

Proof. The estimate (3.7) implies according to Theorem 3.1.2 that equicontinuity condition is satisfied. Since boundedness is immediately implied by equicontinuity and boundedness at zero, to prove precompactness on a dense subset it remains to show by Lemma 3.1.4 that for given $\epsilon > 0$

$$P\left(\forall_{t\in\mathbb{Q}\cap(0,\infty)}\ \forall_{\delta\in(0,\infty)\cap\mathbb{Q}}\ \exists_{N_0(t,\delta)} : \sum_{i=N_0(t,\delta)+1}^{\infty} \|A_i^n(t)\|_{\mathbb{H}}^8 w(i) < \delta\right) > 1-\epsilon. \tag{3.12}$$

For any $\epsilon > 0$ and given fixed t and δ application of Chebyshev inequality in conjunction with the estimate (3.8) yields

$$P\left(\sum_{i=N_0(t,\delta)+1}^{\infty} \|A_i^n(t)\|_{\mathbb{H}}^8 w(i) < \delta\right) = 1 - P\left(\sum_{i=N_0(t,\delta)+1}^{\infty} \|A_i^n(t)\|_{\mathbb{H}}^8 w(i) \geq \delta\right)$$
$$\geq 1 - \frac{E\sum_{i=N_0(t,\delta)+1}^{\infty} \|A_i^n(t)\|_{\mathbb{H}}^8 w(i)}{\delta} > 1-\epsilon.$$

Considering we have only countably many δ's and t's, countable additivity of probability measure gives that (3.12) is indeed fulfilled. □

3.1.2 Solution to the Martingale problem

Now we show that weak limit of sequence $\{P\circ(\tilde{A}^n)^{-1}\}$ can be used to construct martingale solution to the operator (3.4).
We let $A_t(w) = w(t)$, $w \in \Omega$ to be the canonical process on $\Omega = C([0,\infty), S)$ with σ-algebra $\mathcal{F} = \sigma(w(s), s \geq 0)$, $\mathcal{F}_t = \sigma(w(s), 0 \leq s \leq t)$ denotes the usual filtration. We further introduce spaces $\Omega_n = C([0,\infty),(\mathbb{R}^3)^{\Lambda_n})$, $B_t^n(\omega_n) = \omega_n(t)$ the canonical process on Ω_n and the mappings

$$\chi_n : (\mathbb{R}^3)^{\Lambda_n} \to S,\ \chi_n(a_1,\ldots,a_N) = (a_1,\ldots,a_N,0_{i\in\mathbb{Z}^d\setminus\Lambda_n}) \tag{3.13}$$
$$\psi_n : \Omega_n \to \Omega,\ \omega_n \to [t \to (\omega_n(t), 0_{i\in\mathbb{Z}^d\setminus\Lambda_n})]. \tag{3.14}$$

For given $a \in S$ we denote $A^{n,a}$ and $\tilde{A}^{n,a}$ the processes constructed in previous section to accentuate their dependence on a, i. e. $A^{n,a}$ is the solution to SDE with generator extending the (3.5), $A^n(0) = \pi_{\Lambda_n}(a)$ and $\tilde{A}^n = (A^n, 0_{i\in\mathbb{Z}^d\setminus\Lambda_n})$. In addition we denote by P^a the weak limit of measures $P \circ (\tilde{A}^{n,a})^{-1}$, that we just proved in Corollary 3.1.6 to exist. To simplify the notation we denote $\tilde{P}^a_n = P \circ (\tilde{A}^{n,a})^{-1}$ and $P^a_n = P \circ (A^{n,a})^{-1}$. The matching expectations to these three measures will be denoted $E^a, \tilde{E}^a_n$, respectively E^a_n . Notice that $\tilde{P}^a_n = P^a_n \circ \psi_n^{-1}$, as following calculation reveals : for $C \in \mathcal{F}$

$$\tilde{P}^a_n(C) = P(\tilde{A}^{n,a}(\cdot) \in C) = P((A^{n,a}, 0)(\cdot) \in C) = P(\psi_n(A^{n,a}) \in C)$$

$$= P^a_n \circ \psi_n^{-1}(C).$$

We introduce two families of cylindrical functions. We say that $f \in C_c^{2,Cyl}(S)$, if there exists $\Phi_f \subset\subset \mathbb{Z}^d$ such that there is $g \in C_c^2((\mathbb{R}^3)^{\Phi_f}, \mathbb{R})$ and $f(a) = g(\pi_{\Phi_f}(a))$, analogically $f \in C^{2,Cyl}(S)$, if such $g \in C^2((\mathbb{R}^3)^{\Phi_f}, \mathbb{R})$. With this notation we arrive at the following theorem.

Theorem 3.1.7 (Existence of solution to the martingale problem)**.** *Let $a \in S$. Then there exists probability measure P^a on Ω such that :*

$$P(A_0 = a) = 1 \tag{3.15}$$

$$f(A_t) - f(A_0) - \int_s^t Lf(A_u)du \tag{3.16}$$

is $\mathcal{F}_t$-martingale under P^a for any $f \in C_c^{2,Cyl}(S)$ and $\mathcal{F}_t$-local martingale under P^a for any $f \in C^{2,Cyl}(S)$.

Proof. Define P^a as above, so that we have $\tilde{P}^a_n \xrightarrow{w} P^a$. Then with the aid of Portmanteau theorem A.2.3

$$P(A_0 = a) = 1 - \sum_k P^a(\|A_0 - a\|_S > \frac{1}{k})$$

$$\geq 1 - \sum_k \liminf_n P\left(\|\tilde{A}^{n,a}(0) - a\|_S > \frac{1}{k}\right),$$

since by construction $\liminf_n P\left(\|\tilde{A}^{n,a}(0) - a\|_S > 1/k\right) = 0$, we see that (3.15) is indeed satisfied. Let $f \in C_c^{2,Cyl}(S)$ be given. To prove that (3.16) is martingale it suffices to show (following standard procedure in measure theory - see[34] Lemma 3.1) that for arbitrary $G \in C(C([0,s],S),[0,1]), s < t$

$$E^a \left[\left(f(A_t) - f(A_s) - \int_s^t Lf(A_u)du \right) G(\omega.) \right] = 0. \tag{3.17}$$

By weak convergence $\tilde{P}_n^a \xrightarrow{w} P^a$ the formula in (3.17) is a limit of

$$\tilde{E}_n^a \left[\left(f(A_t) - f(A_s) - \int_s^t Lf(A_u)du \right) G(\omega.) \right]. \tag{3.18}$$

We compute

$$\begin{aligned} &\tilde{E}_n^a f(A_t(\omega)) = \tilde{E}_n^a f(\omega_t) = E_n^a f([\psi_n \omega_n]_t) = E_n^a (f \circ \chi_n)(B_t^n(\omega_n)) \\ &\tilde{E}_n^a G(\omega.) = E_n^a G((\psi_n \omega_n).) = E_n^a (G \circ \psi_n)((\omega_n).). \end{aligned} \tag{3.19}$$

Since f is cylindrical the operator L acting on f in fact reduces to L^f, i. e. the operator

$$L^f = \sum_{i \in \Phi_f} \mathcal{L}_{\lambda_i} + q_{x_i} X_i + q_{y_i} Y_i.$$

Consider that for n large enough every point from Φ_f has all neighbours in Λ_n and hence L^f equals to L_n on Φ_f, where L_n is the operator corresponding to A^n as defined in (3.5). Then we adjust

$$\tilde{E}_n^a \int_s^t Lf(A_u) = E_n^a \int_s^t L^f f([\psi_n \omega_n]_u) = E_n^a \int_s^t L^f f(\chi_n(B_u^n(\omega_n))$$

$$= E_n^a \int_s^t L_n(f \circ \chi_n)(B_u^n(\omega_n)).$$

Altogether we found out that (3.18) is equal to

$$E_n^a \left[\left((f \circ \chi_n)(B_t^n) - (f \circ \chi_n)(B_s^n) - \int_s^t L_n(f \circ \chi_n)(B_u^n)du \right) (G \circ \psi_n)((\omega_n).) \right],$$

but since we know that P_n^a solves the martingale problem for L_n on Ω_n, this expression equals to zero and therefore also (3.17) is zero.
The argument, why (3.16) for $f \in C^{2,Cyl}(S)$ is local martingale, is the same

as in finite dimension thanks to the cylindricity assumption. $\square$

3.2 Uniqueness of approximating procedure

In the previous section we showed that our approximation scheme is tight, from which we derived the existence of solution to the martingale problem. It is showed here that the limit point is unique. So although we do not establish uniqueness of martingale problem, we are able to exploit uniqueness of limit point to prove that our solution to martingale problem is indeed Markov process. The proof of uniqueness is based on technical result about behaviour of different approximations schemes in fixed region $\Lambda \subset \mathbb{Z}^d$.

To make the calculations as simple as possible (although still far from trivial) we distinguish specific approximation scheme related to the size of our interactions, namely the boxes Ξ_n. Recall that $0 < r < \infty$ is the parameter of length of interactions for the functions q's and we impose on the interactions additional assumption **(H2)**. This assumption ensures that the equation for A^n has globally Lipschitz drift. More precisely we need the following observation. Remind that the i-th coefficients of equation whose solution is A^n are independent of n provided that $\Pi_i \subset \Lambda_n$, see (3.6) for more detailed discussion there.

Lemma 3.2.1. *Let $\Lambda_n \supset \Xi_{k+1}$ and we denote $b_k = (b_1, \ldots, b_K)$ (notice this does not depend on n, since we assume $\Lambda_n \supset \Xi_{k+1}$) the first $K = |\Xi_k|$ coordinates of drift for the equation*

$$dA^n = b^n(A^n)dt + \sigma^n(A^n)dW_t.$$

For the element $c^n \in (\mathbb{R}^3)^{\Lambda_n}$ we denote $c^n_k = (c^n_{1,x}, \ldots, c^n_{K,z})$. Then there exists constant $J > 0$ s. t.

$$\|b_k(a^n) - b_k(d^n)\|^2_{(\mathbb{R}^3)^{\Xi_k}} \leq J\|a^n_{k+1} - d^n_{k+1}\|^2_{(\mathbb{R}^3)^{\Xi_{k+1}}}, \ \forall a^n, d^n \in (\mathbb{R}^3)^{\Lambda_n}. \quad (3.20)$$

J is independent of k, n.

Proof. The proof is straightforward and follows from assumptions **(H1), (H2), (H3)**. The terms in the drift that complicate Lipschitz condition - and force us to use $k+1$ in (3.20) - are the ones containing $q.$'s, since they depend on all nearest $(2r+1)^d$ neighbours. As an example, how one obtains (3.20) in these cases, we handle using the notation just introduced the term $q_{i,y}(\cdot)a^n_{i,x}$. Because of the finite range of our interactons $q_{i,x}(\cdot)a^n_{i,x}$ is a smooth function of $(2r+1)^d$ variables for $i, 1 \leq i \leq K$, hence application of mean value theorem together with **(H2)** yields

$$\begin{aligned}&\big(q_{i,y}(a^n)a^n_{i,x} - q_{i,y}(d^n)d^n_{i,x}\big)^2 \leq \|\nabla q_{i,y}(\cdot)\cdot_{i,x}\|_\infty \|a^n_{\Pi_i} - d^n_{\Pi_i}\|_{(\mathbb{R}^3)^{(2r+1)^d}} \\ &\leq C\|a^n_{\Pi_i} - d^n_{\Pi_i}\|_{(\mathbb{R}^3)^{(2r+1)^d}}.\end{aligned}$$

We then take into account that every point $i \in \mathbb{Z}^d$ has the same finite fixed amount of neighbours. Hence handling the other terms in the obvious way, we indeed arrive at the existence of some $L > 0$ such that

$$\|b_k(a^n) - b_k(d^n)\|^2_{(\mathbb{R}^3)^{\Xi_k}} \leq L\|a^n_{k+1} - d^n_{k+1}\|^2_{(\mathbb{R}^3)^{\Xi_{k+1}}},\ \forall a^n, d^n \in (\mathbb{R}^3)^{\Lambda_n}.$$

□

We need to restrict our class of starting points $a \in S$, so that the space includes only configurations that does not grow too fast, i. e. we require **(H6)**. The key to proofs in this section are two technical Lemmas about behaviour of solutions A^n to the SDE's related to the operator L_n. If we take some fixed given set $\Gamma \subset \mathbb{Z}^d$ and two supersets $\Gamma_n, \Gamma_k \supset \Gamma$, such that we have corresponding solutions A^n, A^k of SDE's on $(\mathbb{R}^3)^{\Gamma_n}$ resp. $(\mathbb{R}^3)^{\Gamma_k}$, then we cannot claim that $(A^n_i)_{i\in\Gamma}$ and $(A^k_i)_{i\in\Gamma}$ have the same distribution, because we have to redefine the interaction functions at the boundary of the sets Γ_n, Γ_k, and hence $(A^n_i)_{i\in\Gamma}$ and $(A^k_i)_{i\in\Gamma}$ differ as they depend on all A^n resp. A^k via interactions. Therefore we can never have precise equality, even though the coefficients of equations on $(\mathbb{R}^3)^\Gamma$ will be the same for both A^n and A^k, once both Γ_n and Γ_k includes all neighbours of Γ. Nevertheless one would intuitively expect, that the further we are from boundary, the smaller the effect of redefinition should be on Γ. Next Lemma formalizes

and justifies this intuition. Then we can also interpret technical assumption **(H6)** by saying, that the effect of redefining at the boundary will be small, provided we do not start from very fast growing initial configuration.
For the rest of the section we assume conditions **(H1) - (H4), (H6)**.

Lemma 3.2.2. *Let $a \in S$ and Ξ_k be defined as above. Suppose we have two exhausting sequences $\{\Lambda_l\}, \{\Lambda_m\}$ of $\mathbb{Z}^d$ and correspondingly two sequences of processes $\{A^{m,a}\}, \{A^{l,a}\}$. We denote by $A_k^{m,a}$ the part of $A^{m,a}$ on $(\mathbb{R}^3)^{\Xi_k}$, i. e. $A_k^{n,a} = (A_{1,x}^n, \dots, A_{K,z}^n)$. For any $\epsilon > 0$ and $T > 0$ there exists $N > 0$ such that for any $l, m \geq N$*

$$E \sup_{t\in[0,T]} \|A_k^{l,a}(t) - A_k^{m,a}(t)\|^2_{(\mathbb{R}^3)^{\Xi_k}} \leq \epsilon. \tag{3.21}$$

Proof. Not to overload already heavy notation, we will be little ambiguous and write $a_k = (a_{1,x}, \dots, a_{K,z})$ for the restriction of a to $(\mathbb{R}^3)^{\Xi_k}$, while we also denote $a_j = (a_{j,x}, a_{j,y}, a_{j,z})$ when $j \in \mathbb{Z}^d$. Also when dealing with the norms on spaces $\mathbb{R}^n$ for different n we omit the index in the norm, as it should not lead to confusion and instead enhance readability. Using the Lemma 3.2.1 and routine calculations for SDE's we compute

$$\begin{aligned}
&E \sup_{t\in[0,T]} \|A_k^{l,a}(t) - A_k^{m,a}(t)\|^2 \\
&\lesssim E \sup_{t\in[0,T]} \left(\left\| \int_0^t b_k(A^{l,a}) - b_k(A^{m,a}) ds \right\|^2 + \left\| \int_0^t \sigma_k(A^{l,a}) - \sigma_k(A^{m,a}) dW_s \right\|^2 \right) \\
&\lesssim T \left(E \int_0^T \|b_k(A^{l,a}) - b_k(A^{m,a})\|^2 ds + E \int_0^T \|\sigma_k(A^{l,a}) - \sigma_k(A^{m,a})\|^2 ds \right) \\
&\lesssim T \left(\int_0^T E\|A_{k+1}^{l,a}(t_1) - A_{k+1}^{m,a}(t_1)\|^2 dt_1 \right).
\end{aligned}$$

Therefore we obtained the existence of constant $C > 0$ such that

$$E \sup_{t\in[0,T]} \|A_k^{l,a}(t) - A_k^{m,a}(t)\|^2 \leq CT \int_0^T E\|A_{k+1}^{l,a}(t_1) - A_{k+1}^{m,a}(t_1)\|^2 dt_1.$$

Assuming l, m are large enough so we can repeat the procedure above, we get

$$E\|A^{l,a}_{k+1}(t_1) - A^{m,a}_{k+1}(t_1)\|^2 \leq Ct_1 \int_0^{t_1} E\|A^{l,a}_{k+2}(t_2) - A^{m,a}_{k+2}(t_2)\|^2 dt_2$$

$$\cdots \leq C^{n-1} t_1 \int_0^{t_1} t_2 \int_0^{t_2} \cdots \int_0^{t_{n-1}} E\|A^{l,a}_{k+n}(t_n) - A^{m,a}_{k+n}(t_n)\|^2 dt_n \dots dt_1. \tag{3.22}$$

Thanks to the Linear growth of coefficients of our SDE (3.9), there is some $K_T > 0$ such that

$$E\|A^{l,a}_{k+n}(t_n) - A^{m,a}_{k+n}(t_n)\|^2 \leq K_T(1 + \|a_{n+k}\|^2).$$

Using this and then calculating the iterated integrals, we obtain from (3.22) the estimate

$$E \sup_{t\in[0,T]} \|A^{l,a}_k(t) - A^{m,a}_k(t)\|^2 \leq \frac{(CT^2)^n}{(2n-1)!!} K_T(1 + \|a_{n+k}\|^2),$$

where $(2n-1)!! = (2n-1)\cdot(2n-3)\cdots 3\cdot 1$ denotes the odd (double) factorial. Using the trivial bound

$$\|a_{n+k}\|^2_{(\mathbb{R}^3)^{\Xi_{n+k}}} \leq \sum_{j=1}^{(2(n+k)r+1)^d} 3 + \|a_j\|^8_{\mathbb{H}},$$

we need to prove only

$$\lim_{n\to\infty} \frac{L^n}{n!} \sum_{j=1}^{(2(n+k)r+1)^d} (1 + \|a_j\|^8_{\mathbb{H}}) = 0$$

for arbitrary constant $L > 0$. Clearly it suffices to show

$$\lim_n \frac{\sum_{j=1}^{(2(n+k)r+1)^d} \|a_j\|^8_{\mathbb{H}}}{n!^{1-\frac{\delta}{2}}} = 0, \tag{3.23}$$

where δ is from the assumption **(H5)**. We apply **(H5)** and

$$\|a_j\|^8_{\mathbb{H}} u(j) \leq \|a\|^8_S,$$

to conclude that

$$\lim_n \frac{\sum_{j\in\Xi_{n+k+1}\setminus\Xi_{n+k}} \|a_j\|_{\mathbb{H}}^8}{n!^{1-\frac{\delta}{2}}((n+1)^{1-\frac{\delta}{2}}-1)}$$

$$\leq \frac{\|a\|_S^8}{K} \lim_n \frac{(2(n+k+1)r+1)^d - (2(n+k)r+1)^d)(n+k+1)!^{1-\delta}}{n!^{1-\frac{\delta}{2}}} = 0. \tag{3.24}$$

The fact that (3.24) implies (3.23) is well known as Stolz - Cesàro Theorem A.2.4. □

Lemma 3.2.3. *Let $k \in \mathbb{N}$, $a \in S$ and $t > 0$ be given. Let $A^{m,a}$ be approximating sequence defined with respect to exhausting boxes Ξ_m. For any $\epsilon > 0$ there exists $\eta > 0$ such that $\forall m \geq k$*

$$\|b-a\|_S < \eta \Longrightarrow E\|A_k^{m,a}(t) - A_k^{m,b}(t)\|^2 < \epsilon. \tag{3.25}$$

Proof. Since we know that our SDE has continuous dependence on initial condition, the Lemma is nontrivial only for infinite number of m and hence we concentrate in our computations on large m. Again for simplification we will not write the index to the norms through computations. Similarly to the last Lemma we get for some constants $C > 0$ and $K_t > 0$ after n steps (to make last sum meaningful let us formally define $(-1)!! = 1$)

$$E\|A_k^{m,a}(t) - A_k^{m,b}(t)\|^2 \leq C\|a_k - b_k\|^2 + Ct\int_0^t E\|A_{k+1}^{m,a}(t_1) - A_{k+2}^{m,b}(t_1)\|^2 dt_1$$

$$\leq C\|a_k - b_k\|^2 + C^2t^2\|a_{k+1} - b_{k+1}\|^2$$

$$+ Ct\int_0^t Ct_1 \int_0^t E\|A_{k+2}^{m,a}(t_2) - A_{k+2}^{m,b}(t_2)\|^2 dt_2 dt_1$$

$$\leq C\|a_k - b_k\|^2 + C^2t^2\|a_{k+1} - b_{k+1}\|^2 + \cdots + \frac{C^n t^{2n-2}}{(2n-3)!!}\|a_{k+n-1} - b_{k+n-1}\|^2$$

$$+ E \sup_{0\leq s\leq t} \|A_{k+n}^{m,a}(s) - A_{k+n}^{m,b}(s)\|^2 \frac{(Ct^2)^n}{(2n-1)!!}$$

$$\leq \sum_{j=1}^n \frac{C^j t^{2j-2}\|a_{k+j-1} - b_{k+j-1}\|^2}{(2j-3)!!} + K_t \left(\sum_{i\in\Xi_{k+n}} 3 + \|a_i\|_{\mathbb{H}}^8 + \|b_i\|_{\mathbb{H}}^8 \right) \frac{(Ct^2)^n}{(2n-1)!!}.$$

Same calculations like in Lemma 3.2.2 together with Stolz - Cesàro Theorem gives

$$\lim_{n\to\infty} K_t \left(\sum_{i\in\Lambda_{k+n}} 3 + \|a_i\|_{\mathbb{H}}^8 + \|b_i\|_{\mathbb{H}}^8 \right) \frac{(Ct^2)^n}{(2n-1)!!} = 0. \tag{3.26}$$

Because

$$\lim_{n\to\infty} \frac{C^n t^{2n-2} n^l}{((2n-1)!!)^{\frac{\delta}{2}}} = 0$$

for $l > 1$, we obtain using previously established convergence results that

$$\sum_{j=1}^{\infty} \frac{C^j t^{2j-2} \|a_{k+j-1} - b_{k+j-1}\|^2}{(2j-3)!!} < +\infty. \tag{3.27}$$

Therefore combining (3.26) and (3.27) for given $\epsilon > 0$ we can choose $N \in \mathbb{N}$ such that

$$\begin{aligned} &\sum_{j=N}^{\infty} \frac{C^j t^{2j-2} \|a_{k+j-1} - b_{k+j-1}\|^2}{(2j-3)!!} \\ &+ \sup_{j\geq N} K_t \left(\sum_{i\in\Xi_{k+j}} 3 + \|a_i\|_{\mathbb{H}}^8 + \|b_i\|_{\mathbb{H}}^8 \right) \frac{(Ct^2)^j}{(2j-1)!!} < \frac{\epsilon}{2}. \end{aligned}$$

For the first $N-1$ terms we can choose $\eta > 0$ in (3.25) thanks to the continuous dependence on parameters for the $A^{m,a}$ in such way that

$$\begin{aligned} &\sum_{j=1}^{N-1} \frac{C^j t^{2j-2} \|a_{k+j-1} - b_{k+j-1}\|^2}{(2j-3)!!} \\ &+ \sup_{j\leq N-1} E \sup_{0\leq s\leq t} \|A_{k+j}^{m,a}(s) - A_{k+j}^{m,b}(s)\|^2 \frac{(Ct^2)^j}{(2j-1)!!} < \frac{\epsilon}{2}, \end{aligned}$$

and the Lemma is established. □

The first crucial property that follows from Lemma 3.2.2 is independence of the limit measure P^a on the choice of convergent subsequence. By the well known properties of weak convergence this implies that the sequence $\{\tilde{P}_n^a\}$

itself weakly converges. In addition this limit doesn't depend on the choice of approximating sequence Λ_n.

Theorem 3.2.4. *Let $\tilde{A}^{m,a}, \tilde{A}^{n,a}$ be the sequences of approximating processes on Ω, $a \in S$. Then there exists probability measure P^a on Ω such that*

$$\lim_{m\to\infty} P \circ (\tilde{A}^{m,a})^{-1} = \lim_{l\to\infty} P \circ (\tilde{A}^{l,a})^{-1} = P^a.$$

Proof. By Corollary 3.1.6 we know that any two such sequences has weakly convergent subsequence. So it remains to show that the limit point is the same for any two weakly convergent subsequences (to simplify notation we call the convergent subsequences again m and l) $\{P\circ(\tilde{A}^{l,a})^{-1}\}, \{P\circ(\tilde{A}^{m,a})^{-1}\}$. To prove this it clearly suffices to show that for any $f \in C_b(\Omega)$

$$\lim_l Ef(\tilde{A}^{l,a}(\cdot)) = \lim_m Ef(\tilde{A}^{m,a}(\cdot)). \tag{3.28}$$

First let $f \in C^{Cyl}_{b,Lip}(\Omega)$, i. e. there exists $k \in \mathbb{N}$ and $g \in C_{b,Lip}(\Omega_{\Xi_k})$ such that $f(\omega) = g((\pi_{\Xi_k}\omega).)$, $\Omega_{\Xi_k} = C([0,\infty), (\mathbb{R}^3)^{\Xi_k})$ and g is Lipschitz, that is there exists constant $L > 0$ s. t.

$$|g((\omega_k).) - g((\tilde{\omega}_k).)| \le \|\omega_k - \tilde{\omega}_k\|_{\Omega_{\Xi_k}} \;\; \forall \omega_k, \tilde{\omega}_k \in \Omega_{\Xi_k}.$$

Then we get for m, l large enough

$$\begin{aligned} &|Ef(\tilde{A}^{l,a}(\cdot)) - Ef(\tilde{A}^{m,a}(\cdot))|^2 = |Eg(A^{l,a}_k(\cdot)) - Eg(A^{m,a}_k(\cdot))|^2 \\ &\le E|g(A^{l,a}_k(\cdot)) - g(A^{m,a}_k(\cdot))|^2 \le E\|A^{l,a}_k(\cdot) - A^{m,a}_k(\cdot)\|^2, \end{aligned}$$

hence Lemma 3.2.2 implies (3.28) holds for $f \in C^{Cyl}_{b,Lip}(\Omega)$.
Next let $f \in C^{Cyl}_b(\Omega)$, then there exists bounded sequence $f_n \in C^{Cyl}_{b,Lip}(\Omega)$ such that $f_n \to f$. Finally for $f \in C_b(\Omega)$ consider cylindrical approximation by $\{f_n\}$, that is $f_n(\omega.) = f((\pi_{\Xi_n}\omega).)$ and the result follows by Lebesgue's dominated convergence theorem. $\square$

3.2.1 Markov property

To translate Lemma 3.2.3 into desired properties, we recall equivalent definition of weak convergence. Its proof follows immediately from Skorokhod representation theorem (see also[72] pp. 168).

Lemma 3.2.5. *Let $\mathcal{S}$ be a Polish space and μ_n, μ probability measures on $\mathcal{S}$. Suppose $\mu_n \xrightarrow{w} \mu$. Let $\{f_n\}, f \in C(P)$ be such that functions $\{f_n\}$ are uniformly bounded and*

$$x_n \to x \text{ in } \mathcal{S} \Longrightarrow f_n(x_n) \to f(x). \tag{3.29}$$

Then $\mu_n f_n \to \mu f$.

We apply this Lemma to show that canonical process on Ω is a genuine Markov process under the measures P^a.

Theorem 3.2.6. *Let $A_t(w)$ be canonical process on $\Omega = C([0,\infty), S)$ and P^a the unique limiting measure produced by Corollary 3.1.6. (A_t, P^a) is then a Markov process.*

Proof. Denote $\mathcal{S}$ the σ-algebra on S. We need to show these two properties

$$\textbf{(I)}\ a \to P^a(A_t \in C) \text{ is measurable for any } C \in \mathcal{S} \tag{3.30}$$

$$\textbf{(II)}\ P^a(A_{s+t} \in C|\mathcal{F}_s) = \phi(A_s),\ \phi(\cdot) = P^{\cdot}(A_t \in B),\ \forall\ C \in \mathcal{S}, 0 \le s \le t. \tag{3.31}$$

To prove (3.30) we show that $a \to E^a f(A(t))$ is continuous function for any $f \in C_{b,Lip}^{Cyl}(S)$, the measurability for general $f \in C_b(S)$ will then follow through same procedure as in Theorem 3.2.4. By the uniqueness just proved, we can consider approximation $\{A^n\}$ living on the boxes Ξ_n. So let $f(a) = g(\pi_{\Xi_k}(a))$, we then calculate

$$|E^a f(A_t) - E^b f(A_t)|^2 = |\lim_n E[f(\tilde{A}^{n,a}(t)) - f(\tilde{A}^{n,b}(t))]|^2$$
$$\le \limsup_n E|g(A_k^{n,a}(t)) - g(A_k^{n,b}(t))|^2 \le \limsup_n \|A_k^{n,a}(t) - A_k^{n,b}(t)\|^2.$$

From Lemma 3.2.3 we derive that this estimate establishes the desired continuity.
To prove (3.31) one strives to establish $\forall f \in C_b(S)$

$$E^a[f(A_{s+t})|\mathcal{F}_s] = E^{A_s} f(A_t). \tag{3.32}$$

If we denote $\varphi(\cdot) = E^{\cdot}[f(A_t)]$, this then means - for any $C \in \mathcal{F}_s$

$$\int_C f(A_{s+t}) dP^a = \int_C \varphi(A_s) dP^a.$$

We consider first $f \in C_{b,Lip}^{Cyl}(S)$, then we know from the first part of the proof that $\varphi(\cdot)$ is continuous. By approximation this reduces to necessity of demonstrating

$$E^a[f(A_{s+t})h(\omega.)] = E^a \varphi(A_s) h(\omega.), \tag{3.33}$$

where h is arbitrary, but fixed continuous bounded $\mathcal{F}_s$ - measurable function. By weak convergence $\tilde{P}_n^a \to P^a$ the left side of (3.33) is a limit of (the same calculations as we made in the proof of Theorem 3.1.7 are hidden there)

$$\tilde{E}_n^a[f(A_{s+t})h(\omega.)] = E_n^a[(f \circ \chi_n)(B_{s+t}^n)(h \circ \psi_n)((\omega_n).)].$$

The finite dimensional result, i. e. the fact that P_n^a solves the martingale problem on Ω_n, tells us that

$$E_n^a[(f \circ \chi_n)(B_{s+t}^n)(h \circ \psi_n)((\omega_n).)] = E_n^a[\varphi^n(\chi_n(B_s^n))(h \circ \psi_n)((\omega_n).)],$$

if $\varphi^n(\chi_n(B_s^n)) = E_n^{\chi_n(B_s^n)}[(f \circ \chi_n)(B_t^n)]$. We observe that

$$\varphi(a) = E^a f(A_t) = \lim_n \tilde{E}_n^a f(A_t) = \lim_n E_n^a[(f \circ \chi_n)(B_t^n)],$$

hence (3.33) will established using Lemma 3.2.5, provided we can prove the implication

$$a_n \to a \text{ in } S \Longrightarrow \tilde{E}_n^{a_n}[f(A_t)] \to E^a[f(A_t)]. \tag{3.34}$$

For given $\epsilon > 0$ we find N from weak convergence such that

$$|E^a[f(A_t)] - \tilde{E}_n^a[f(A_t)]| < \frac{\epsilon}{2} \ \forall n \geq N.$$

Like in the first part we also have the estimate

$$|\tilde{E}_n^{a_n}[f(A_t)] - \tilde{E}_n^a[f(A_t)]|^2 \leq \limsup_n \|A_k^{n,a_n}(t) - A_k^{n,a}(t)\|^2,$$

so Lemma 3.2.3 implies we can find $\delta > 0, \tilde{N} \in \mathbb{N}$ such that

$$\|a - a_n\| < \delta \implies |E^a[f(A_t)] - \tilde{E}_n^{a_n}[f(A_t)]| < \epsilon \ \forall n \geq \tilde{N}.$$

Therefore from Lemma 3.2.5 we conclude that (3.33) holds for $f \in C_{b,Lip}^{Cyl}(S)$. We infer the validity of (3.32) for general $f \in C_b(S)$ by routine approximation procedure. $\square$

This result implies that if we set $P_t(a, C) = P^a\{A_t \in C\}$, then P_t is a genuine transition probability function and $P_t f(a) = E^a f(A(t))$ is the Markov semigroup acting on all $f \in \mathcal{B}_b(S)$ satisfying the Chapman - Kolmogorov equality[12] chap. I.

3.3 Existence of invariant measure for the semigroup

We now derive the tightness of measures $\{\nu_n\}$ and consequently show that any limit point is an invariant measure for the Markov semigroup constructed above.
We need to enlarge our space S to assure that it accommodates invariant measure. The assumption that fits the bill is **(H5)**. Thus for the remainder of the chapter we work under assumptions **(H1) - (H6)**.

Theorem 3.3.1. *The sequence of measures $\{\nu_n\}$ is tight.*

Proof. We want to show that for given $\epsilon > 0$ there is compact set K_ϵ in S such that $\forall n \in \mathbb{N}$ one has $\nu_n(K_\epsilon) \geq 1 - \epsilon$.

Let us recall the estimate (2.4)

$$L_n W_n^2(a_n) \leq -cW_n^2(a_n) + C. \tag{3.35}$$

Remind that $\nu_n = \mu_n \circ \chi_n^{-1}$ (3.13), and μ_n is an invariant measure for process A^n on $(\mathbb{R}^3)^{\Lambda_n}$. Hence we have the equality

$$\mu_n(L_n W_n^2) = 0. \tag{3.36}$$

Clearly

$$\mu_n(L_n W_n^2) = \mu_n(L_n W_n^2 I_{L_n W_n^2 > 0}) + \mu_n(L_n W_n^2 I_{L_n W_n^2 \leq 0}),$$

and from (3.35) it follows that $\mu_n(L_n W_n^2 I_{L_n W_n^2 > 0}) \leq C$, so that

$$\mu_n(L_n W_n^2 I_{L_n W_n^2 \leq 0}) \geq -C. \tag{3.37}$$

Notice that in our notation it holds

$$W_n^2(a_n) = \sum_{i=1}^{N} V_i^2(a_i) v(i) = \sum_{i=1}^{N} v(i) \|a_i\|_{\mathbb{H}}^8.$$

For given $\epsilon > 0$ we define the set K_ϵ as

$$K_\epsilon = \left\{ a \in S : \forall i \in \mathbb{Z}^d \ \|a_i\|_{\mathbb{H}}^8 u(i) \leq u(i) \left(\frac{C+1}{c\epsilon v(i)} + \frac{C}{cv(i)} \right) \right\}.$$

Thanks to the assumption **(H5)** this set is compact in S according to Lemma 3.1.4. We calculate

$$\nu_n(K_\epsilon^C) = \mu_n(\chi_n(K_\epsilon^C)) = \mu_n \left(b \in (\mathbb{R}^3)^{\Lambda_n} : \ \exists i \in \Lambda_n : \|b_i\|_{\mathbb{H}}^8 > \frac{C+1}{c\epsilon v(i)} + \frac{C}{cv(i)} \right).$$

Hence for $a_n \in \chi_n(K_\epsilon^C)$ we have

$$L_n W_n^2(a_n) \leq -c \left(\frac{C+1}{c\epsilon v(i)} + \frac{C}{cv(i)} \right) v(i) + C \leq -\frac{C+1}{\epsilon}.$$

Therefore if $\nu_n(K_\epsilon^C) > \epsilon$ holds, we get the contradiction with (3.37), which finishes the proof. $\square$

Theorem 3.3.2. *There exists an invariant measure for the Markov process (A, P^a) from Theorem 3.2.6.*

Proof. We fix some weakly convergent sequence of measures $\{\nu_n\}$ and its limit point ν. To show that ν is invariant, we prove that for any $f \in C_b(S)$

$$\int_S P_t f(a) d\nu(a) = \int_S f(a) d\nu(a). \tag{3.38}$$

We show (3.38) holds for $f \in C^{Cyl}_{b,Lip}(S)$, the general case follows by approximation as before. Recall that $\nu_n = \mu_n \circ \chi_n^{-1}$ and that μ_n is the invariant measure for process A^n on $(\mathbb{R}^3)^{\Xi_n}$, so that the equality

$$\int_{(\mathbb{R}^3)^{\Xi_n}} E_n^{i_n(e_n)} h(B_t^n) d\mu_n(e_n) = \int_{(\mathbb{R}^3)^{\Xi_n}} h(e_n) d\mu_n(e_n) \quad \forall h \in C_b((\mathbb{R}^3)^{\Xi_n})$$

holds. Arguing like in the proof of (3.31), we may use Lemma 3.2.5 to prove the equality

$$\lim_n \int_S \tilde{E}_n^a f(A_t) d\nu_n(a) = \int_S E^a f(A_t) \nu(a).$$

Remembering the calculations (3.19) we compute

$$\begin{aligned}
&\int_S f(a) d\nu(a) = \lim_n \int_S f(a) d\nu_n(a) = \lim_n \int_{(\mathbb{R}^3)^{\Xi_n}} (f \circ \chi_n)(a_n) d\mu_n(a_n) \\
&= \lim_n \int_{(\mathbb{R}^3)^{\Xi_n}} E_n^{\chi_n(a_n)} (f \circ \chi_n)(B_t^n) d\mu_n(a_n) = \lim_n \int_S E_n^a (f \circ \chi_n)(B_t^n) d\nu_n(a) \\
&= \lim_n \int_S \tilde{E}_n^a f(A_t) d\nu_n(a) = \int_S E^a f(A_t) \nu(a) = \int_S P_t f(a) d\nu(a),
\end{aligned}$$

what we wanted to show. $\square$

3.4 Examples of other operators

We list some other relevant examples, that can be handled using our strategy without any additional difficulty :

- Of course the elliptic case lies naturally within our framework. Take standard Euclidean space $\mathbb{R}^3$ with standard Laplacian Δ, $D = x\partial_x + y\partial_y + z\partial_z$, $X = \partial_x$ (etc. for Y, Z), $\mathcal{L}_\lambda = \Delta - \lambda D$ and consider operator

$$L = \sum_{i\in\mathbb{Z}^d} \mathcal{L}_{\lambda_i} + q_{i,x}X_i + q_{i,y}Y_i + q_{i,z}Z_i$$

acting on $(\mathbb{R}^3)^{\mathbb{Z}^d}$. Lyapunov function here can be chosen just $x^{2k} + y^{2k} + z^{2k}$, for $k = 2$ we get the same tightness as we had in Corollary 3.1.6.

- The Grushin plane[1] : Take $\mathbb{R}^2$ as the basic space and consider vector fields $X = \partial_x, Y = -x\partial_y$. D is given by $D = x\partial_x + y\partial_y$ and operator

$$L = \sum_{i\in\mathbb{Z}^d} X_i^2 + Y_i^2 - \lambda_i D_i + q_{i,x}X_i + q_{i,y}Y_i$$

on $(\mathbb{R}^2)^{\mathbb{Z}^d}$. For the Lyapunov function works $V = x^{4k} + y^{2k}$, the tightness (3.1.6) works again for $k = 2$. The σ and u in Girsanov theorem to simplify the control problem can be chosen in the following way

$$\sigma = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2}x \end{pmatrix} \quad u = \begin{pmatrix} \frac{q_x}{\sqrt{2}} \\ \frac{q_y}{x} \end{pmatrix}.$$

Then we have

$$\sigma u = b - \tilde{b} = (-\lambda x, -\lambda y).$$

- We cannot quite handle the example of Martinet distribution as in[25]. Take $\mathbb{R}^3$ and let $X = \partial_x - y^2\partial_z$, $Y = y\partial_y$. The problem that arises lies in the nonlinear term in z-axis. We can not hope for our strategy to be successful, as in the last section definitely linear growth together with strong Lipschitz condition is required. But at least the finite dimensional case is almost conquered by our methods - If one puts $D = x\partial_x + y\partial_y + z\partial_z$ and consider

$$L = X^2 + Y^2 - \lambda D + q_xX + q_yY$$

as operator on $\mathbb{R}^3$, then the SDE corresponding to this operator has coefficients

$$b = (q_x - \lambda x, q_y - \lambda y, -\lambda z - q_x y^2),\ \sigma = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2}y^2 \end{pmatrix}.$$

Due to nonlinearities, not even global existence of process is a priori clear. However, if we set $V_k = x^{2k} + y^{6k} + z^{2k}$, we calculate that V_k is the Lyapunov function giving global existence and invariant measure. The smoothness of density holds from Theorem 2.2.1 as well. However to our best knowledge, we are unable to investigate the irreducibility of the process.

In general we can say, that our strategy is successful whenever we can establish finite dimensional results as in Theorem 2.1.2 with Lyapunov function, that will enable us to construct the diffusion using tightness arguments as in chapter three. To finish the strategy with desired results, it is then essential that we can impose on the interaction such constraints that lead to the conditions in Lemma 3.2.1.

4

Variational approach to stochastic partial differential equations

We outline the basic results about so-called variational approach to stochastic differential equations in infinite dimension. This approach was founded by Pardoux in[65] and Krylov and Rozovskii in[45]. The chapter serves reference purpose and is motivated by our use of these solution in the next one. Our presentation is based on books[68],[54], see also[19] for more recent treatment.

4.1 Definition of solution and admissible coefficients

Let H and K be real separable Hilbert spaces, B a reflexive Banach space embedded continuously and densely in H. Upon identifying H with its dual H^* we get a Gelfand triple $B \subseteq H \subseteq B^*$; note that – in this representation – the restriction of the dual pairing $\langle \cdot, \cdot \rangle_{B^*,B}$ to $H \times B$ coincides with the scalar product $\langle \cdot, \cdot \rangle_H$ in H. Assume that

(A1) *$f\colon \mathbb{R}_{\geq 0} \times B \longrightarrow B^*$ and $\sigma\colon \mathbb{R}_{\geq 0} \times B \longrightarrow \mathscr{L}_2(K,H)$ (for definition of Hilbert-Schmidt operators see (A.1.3)) are Borel functions, μ is a Borel probability measure on H,*

and consider a stochastic evolution equation

$$\mathrm{d}X = f(t,X)\,\mathrm{d}t + \sigma(t,X)\,\mathrm{d}W, \quad X_0 \sim \mu. \tag{4.1}$$

Note that in[54] the authors admit random coefficients, but as we do not need it for our purposes, we stick to deterministic case. Randomness does not bring additional difficulties, except perhaps complicating notation and assumptions. The variational solution to equation (4.1) we shall define as follows.

Definition 4.1.1 (variational solution)**.** A triple $((\Omega, \mathscr{F}, (\mathscr{F}_t), \boldsymbol{P}), W, X)$ is called *a (variational) solution* to the stochastic evolution equation (4.1) provided $(\Omega, \mathscr{F}, (\mathscr{F}_t), \boldsymbol{P})$ is a stochastic basis satisfying the usual conditions on which a standard cylindrical $(\mathscr{F}_t)$-Wiener process W (A.1.2) on K and a B^*-valued $(\mathscr{F}_t)$-progressively measurable process X are defined such that

i) $X(0)_\# \boldsymbol{P} = \mu$ (i.e., the law of $X(0)$ is μ),

ii) there exists an $(\mathscr{F}_t)$-progressively measurable B-valued process $\tilde{X}$ satisfying $\|\tilde{X}\|_B \in L^p_{\mathrm{loc}}(\mathbb{R}_{\geq 0})$ $\boldsymbol{P}$-almost surely for some $p \in (1,\infty)$, $X = \tilde{X}$ $\lambda \otimes \boldsymbol{P}$-almost everywhere on $\mathbb{R}_{\geq 0} \times \Omega$,

iii) $\|f(\cdot, \tilde{X}(\cdot))\|_{B^*}^{p/(p-1)} + \|\sigma(\cdot, \tilde{X}(\cdot))\|^2_{\mathscr{L}_2(K,H)} \in L^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0})$ and

$$X(t) = X(0) + \int_0^t f(s, \tilde{X}(s))\,\mathrm{d}s + \int_0^t \sigma(s, \tilde{X}(s))\,\mathrm{d}W_s \quad \text{in } B^*$$

for all $t \geq 0$ $\boldsymbol{P}$-almost surely.

We follow the presentation in[54], hence we distinguish formally between X and $\tilde{X}$. Note though that $\tilde{X}$ is just progressively measurable version of X in smaller space B. Instead of assuming the existence of $\tilde{X}$, we could have assumed the existence of progressively measurable modification of X with respect to this smaller space B.

It turns out, that we need to impose the following four conditions upon the coefficients.

(G1) (Hemicontinuity) For all $u, v, w \in B$, and $t \in [0, \infty)$ the map

$$\lambda \rightarrow_{B^*} \langle f(t, u + \lambda v), w\rangle_B, \ \lambda \in \mathbb{R}$$

is continuous.

(G2) (Weak monotonicity) There exist $c \in \mathbb{R}$ such that for all $u, v \in B$ and $T > 0$

$$\begin{aligned} &2_{B^*}\langle f(\cdot, u) - f(\cdot, v), u - v\rangle_B + \|\sigma(\cdot, u) - \sigma(\cdot, v)\|^2_{\mathscr{L}_2(K,H)} \\ &\leq c\|u - v\|_H \text{ on } [0, T] \end{aligned}$$

(G3) (Coercivity) There exist $\alpha \in (1, \infty)$, $c_1 \in \mathbb{R}$, $c_2 \in (0, \infty)$ and $g \in L^1_{\text{loc}}(\mathbb{R}_{\geq 0})$ such that for all $v \in B, t \geq 0$

$$2_{B^*}\langle f(t, v), v\rangle_B + \|\sigma(t, v)\|^2_{\mathscr{L}_2} \leq c_1\|v\|^2_H - c_2\|v\|^\alpha_B + g(t).$$

(G4) (Boundedness) There exist $c_3 \in (0, \infty)$ and $h \in L^{\frac{\alpha}{\alpha-1}}_{\text{loc}}(\mathbb{R}_{\geq 0})$ such that for all $v \in V, t \geq 0$

$$\|f(t, v)\|_{B^*} \leq h(t) + c_3\|v\|^{\alpha-1}_B,$$

where α is from (G3).

When investigating, whether particular f and σ satisfy (G1) - (G4), one can focus on f in isolation. Indeed, if f satisfies (G2), (G3) and for all $t \geq 0$ the map $u \rightarrow \sigma(t, u)$ is Lipschitz with constant independent of t, then f and σ satisfy (G2), (G3).
From linearity we also see that if f, σ satisfy (G2), (G3), and $\tilde{f}$ satisfies (G2),(G3), then $f + \tilde{f}, \sigma$ satisfy (G2), (G3).
Likewise, if f and $\tilde{f}$ satisfy (G1), (G4), so does $f + \tilde{f}$.

4.2 Main Existence result and Examples

The basic theorem about existence of variational solution we copy from [54] pp. 91 following pioneering work [45].

Theorem 4.2.1. *Let f, σ be functions as in (A1) satisfying (G1)-(G4) with $X_0 \in L^2(H)$ $\boldsymbol{P}$-almost surely. Then there exists a unique solution X to (4.1) in the sense of Definition 4.1.1. Moreover*

$$E(\sup_{t\in[0,T]} \|X(t)\|_H^2) < \infty.$$

Let us provide some interesting examples of coefficients related to our work in next chapter, that satisfy above described condition (G1) - (G4).

Example 4.2.1. We cover the case, when f is given by Laplace operator Δ, consult [54] Example 4.1.7 for more details. Let $\tilde{D} \subseteq \mathbb{R}^d$ be bounded open set with sufficiently smooth boundary ∂D. We want f to be an extension of Δ to a properly chosen Banach space B so that $f : B \to B^*$ and f is continuous with respect to $\|\cdot\|_B$ and $\|\cdot\|_{B^*}$. Natural choice is the classical Sobolev space $W_0^{1,2}(\tilde{D})$ with Dirichlet boundary condition (A.1.4). It is readily shown that $W_0^{1,2}(\tilde{D})$ is embedded continuously and densely (cf. [16]) in $L^2(\tilde{D})$. Thence upon identifying $L^2(\tilde{D})$ with its dual we get Gelfand triple

$$B = W_0^{1,2}(\tilde{D}) \subset H = L^2(\tilde{D}) \subset B^* = W_0^{1,2}(\tilde{D})^*.$$

To extend Δ with initial domain $C_c^\infty(G)$ to a bounded linear operator

$$f : B \to B^*$$

we note that Δu for $u \in C_c^\infty$ has values in B^*. Using integraton by parts for $u, v \in B$ we obtain

$$\begin{aligned}|{}_{B^*}\langle \Delta u, v\rangle_B| = |\langle \Delta u, v\rangle_H| &= \left| -\int_{\tilde{D}} \langle \nabla u(x), \nabla v(x)\rangle dx \right| \\ &\leq \|\nabla u\|_{L^2} \|\nabla v\|_{L^2} \leq \|\nabla u\|_{L^2} \|v\|_{1,2}.\end{aligned}$$

Therefore for all $u \in C_c^\infty(\tilde{D})$ we get

$$\|\Delta u\|_{B^*} \leq \|\nabla u\|_{L^2} \leq \|u\|_{1,2}. \tag{4.2}$$

Thus Δ with domain C_c^∞ extends uniquely to a bounded linear operator $f : B \to B^*$. We claim that (G1)-(G4) holds for f.

(G1) is obvious from linearity of A. For $u, v \in B$ consider sequences $u_n, v_n \in C_c^\infty(\tilde{D})$ such that $u_n \to u$, $v_n \to v$ in B. Again relying on integration by parts

$$\begin{aligned} {}_{B^*}\langle A(u) - A(v), u - v\rangle_B &= \lim_{n\to\infty} \langle \Delta(u_n - v_n), u_n - v_n\rangle_H \\ &= -\lim_{n\to\infty} \int |\nabla(u_n - v_n)(x)|^2 dx \leq 0, \end{aligned}$$

which shows (G2). Furthermore,

$$\begin{aligned} {}_{B^*} < A(v), v >_B = \lim_{n\to\infty} < \Delta v_n, v_n >_H = -\lim_{n\to\infty} \int |\nabla v_n(x)|^2 dx \\ = -\int |\nabla v(x)|^2 dx \leq (\|v\|_H^2 - \|v\|_{1,2}^2). \end{aligned}$$

So (G3) is satisfied with $\alpha = 2$ and (G4) with $\alpha = 2$ is clear by (4.2).

Example 4.2.2. With B and H as above, we consider the so called p-Laplacian $\Delta_p, p \geq 2$, on a open bounded set $\tilde{D}$,[54] Example 4.1.9. Precisely, for given $u \in B = W_0^{1,2}(\tilde{D})$ define $f : W_0^{1,2}(\tilde{D}) \to W_0^{1,2}(\tilde{D})^*$ by

$${}_{B^*}\langle f(u), v)\rangle_V := -\int_{\tilde{D}} |\nabla u(x)|^{p-2} \langle \nabla u(x), \nabla v(x)\rangle dx \ \forall v \in B. \tag{4.3}$$

Notice that for $p = 2$ we get ordinary Laplacian Δ. To show that $f : B \to B^*$ is well-defined operator, one simply uses Hölder's inequality to get

$${}_{B^*}\langle f(u), v\rangle_B \leq \|u\|_{1,p}^{p-1} \|v\|_{1,p},$$

so $f(u)$ is well-defined element of $W_0^{1,p}(\tilde{D})^*$ and we have estimate

$$\|f(u)\|_{B^*} \leq \|u\|_B^{p-1}.$$

This estimate clearly implies (G4) with $\alpha = p$. (G1) and (G2) follows by rather routine calculations ([54] pp. 82). (G3) is more intriguing as the proof uses Poincaré's inequality A.2.5. For $u \in B$ we compute

$$_{B*}\langle f(u), u\rangle_B = -\int |\nabla u(x)|^p dx \le -\frac{1 \wedge C}{2}\|u\|_{1,p}^p,$$

where C is from Poincaré's inequality. Therefore (G3) holds with $\alpha = p$ and $c_1 = 0$.

Example 4.2.3. We present porous medium operator here, more details may be found e.g. in[54] Example 4.1.11 and[69]. Let again $\tilde{D}$ be open bounded set with smooth boundary and $p \ge 2$, we set $B = L^p(\tilde{D})$, $H = W_0^{1,2}(\tilde{D})^*$. For $u \in W_0^{1,2}(\tilde{D})$ we define

$$\|u\|_{W_0^{1,2}} := \left(\int_{\tilde{D}} |\nabla u(x)|^2 dx\right)^{\frac{1}{2}}.$$

Now Poincaré's inequality A.2.5 implies that this is equivalent norm with $\|\cdot\|_{1,2}$, so one can consider $W_0^{1,2}(\tilde{D})$ with this norm and corresponding scalar product

$$\langle u, v\rangle_{W_0^{1,2}} = \int \langle \nabla u(x), \nabla v(x)\rangle dx, \;\; u, v \in W_0^{1,2}(\tilde{D}).$$

Since $W_0^{1,2}(\tilde{D}) \subset (L^{\frac{p}{p-1}})$ continuously and densely, we also have $L^p(\tilde{D}) \subset W_0^{1,2}(\tilde{D})^*$ continuously and densely, hence indeed $B \subset H \subset B^*$ is Gelfand triple. Space $H^* = W_0^{1,2}(\tilde{D})$ can be identified with H via Riesz isomorphism. Recall the definition of Laplacian Δ in this context (4.3). We note however that the dualities appearing in this example must be handled with some care. In particular consider the following Lemma ([54] pp. 85 for proof).

Lemma 4.2.2. *The map*

$$\Delta : W_0^{1,2}(\tilde{D}) \to (L^p(\tilde{D}))^*$$

extends to a linear isomorphic isometry

$$\Delta L^{\frac{p}{p-1}}(\tilde{D}) \to (L^p(\tilde{D}))^* = B^*$$

and

$$_{B^*}\langle \Delta u, v\rangle_B = -_{L^{\frac{p}{p-1}}}\langle u, v\rangle_{L^p} = -\int u(x)v(x)dx$$

$$\forall u \in L^{\frac{p}{p-1}}(\tilde{D}), v \in L^p(\tilde{D}).$$

Therefore we have surprising dualization between L^p and $(L^p)^*$, since

$$(L^p)^* = \Delta(L^{\frac{p}{p-1}}) \neq L^{\frac{p}{p-1}}.$$

To define porous medium operator, we let $\Psi : \mathbb{R} \to \mathbb{R}$ be a function with the following properties:

(Ψ1) Ψ is continuous

(Ψ2)

$$(t-s)(\Psi(t) - \Psi(s)) \geq 0 \;\forall s, t \in \mathbb{R}$$

(Ψ3) There exist $p \geq 2$, $a > 0$, $c \geq 0$ such that for all $s \in \mathbb{R}$

$$s\Psi(s) \geq a|s|^p - c.$$

(Ψ4) There exist $c_3, c_4 > 0$ such that for all $s \in \mathbb{R}$

$$|\Psi(s)| \leq c_4 + c_3|s|^{p-1},$$

where p is from (Ψ3).

Notice that (Ψ4) implies that

$$\Psi(v) \in L^{\frac{p}{p-1}}(\tilde{D}) \;\forall v \in L^p(D).$$

Following this, we define porous medium operator $A : B \to B^*$ by setting

$$A(u) := \Delta\Psi(u),\; u \in L^p(\tilde{D}).$$

One routinely checks that (G1)-(G4) are now satisfied for such operator A. Typical example of function satisfying (Ψ1)-(Ψ4) provides porous medium equation, i. e. for $p \in (2, \infty)$ we take function $\Psi(s) = s|s|^{p-2}$.

5

Continuous-time stochastic approximation in infinite dimensions

Stochastic approximation was originally introduced as a procedure for sequentially finding a zero or an extremum point of a function which can be observed only with a random measurement error; it has found many applications e.g. to recursive estimation, adaptive control or learning algorithms, see the books[9,10,15,18,46] or[48] for a thorough information about the stochastic approximation methods. The seminal Robbins-Monro procedure may be roughly described as follows: Let $R\colon \mathbb{R} \to \mathbb{R}$ be a function which is known to have a unique root x_0 but the observation of $R(x)$ at time $k \in \mathbb{N}$ is corrupted by a noise $e_k(x)$. Let $\alpha_n > 0$ be such that

$$\sum_{n=0}^{\infty} \alpha_n = \infty, \quad \sum_{n=0}^{\infty} \alpha_n^2 < \infty$$

and set

$$Y_{n+1} = Y_n + \alpha_n \big(R(Y_n) + e_{n+1}(Y_n) \big).$$

Then under suitable assumptions upon the function R and the random variables $e_k(x)$ it may be shown that $Y_n \to x_0$ almost surely as $n \to \infty$. M. B. Nevel'son and R. Z. Khas'minskiĭ in their book[61] studied a continuous-time version of stochastic approximation. In particular, they introduced a continuous-time analogue of the Robbins-Monro procedure: Consider a stochastic differential equation

$$\mathrm{d}X = \alpha(t)\big(R(X)\,\mathrm{d}t + \sigma(t,X)\,\mathrm{d}W\big), \quad X_0 = x, \tag{5.1}$$

where W is a Wiener process and α is a strictly positive function in $L^2(\mathbb{R}_{\geq 0}) \setminus L^1(\mathbb{R}_{\geq 0})$. Sufficient conditions for X_t to converge to the zero set of R almost surely as $t \to \infty$ were found in terms of existence of a suitable Lyapunov function for (5.1). One may consult the book[43] or the papers[66,17,49,50,51,52] for further results in this direction. Due to powerful tools from stochastic analysis, proofs in the continuous-time case may be presented in a very lucid way (cf. also[17] for a discussion of this point).

Our aim in this chapter is to extend the stochastic analysis approach, in the form proposed by Nevel'son and Khas'minskiĭ, to infinite-dimensional Hilbert spaces. Several results on discrete-time stochastic approximation in infinite-dimensional spaces are available, cf. e.g.[6,29,47,62,79,80], but the only paper using infinite-dimensional stochastic analysis to study stochastic approximation we are aware of is[5] § 4. However,[5] treats stochastic delay equations, whilst we are interested in stochastic partial differential equations. We confine ourselves to procedures of the Robbins-Monro type in the case of a unique root, since we see our task in indicating how the ideas from[61] may be combined with techniques from the theory of stochastic evolution equations, not in obtaining the strongest possible results. A typical example we can cover is the following: Consider a nonlinear elliptic equation

$$\Delta u + r(u) = f \text{ in } D, \quad u = 0 \text{ on } \partial D, \tag{5.2}$$

where $\tilde{D} \subseteq \mathbb{R}^d$ is a bounded domain with a smooth boundary $\partial\tilde{D}$, and a stochastic parabolic equation

$$\mathrm{d}X = \alpha(t)\big(\Delta X + r(X) - f\big)\,\mathrm{d}t + \alpha(t)\sigma(t,X)\,\mathrm{d}W, \quad X_{|\mathbb{R}_{>0}\times\partial\tilde{D}} = 0,\ X_0 = y \tag{5.3}$$

in $L^2(D)$, driven by an infinite-dimensional Wiener process W, where $\alpha \in L^2(\mathbb{R}_{\geq 0}) \setminus L^1(\mathbb{R}_{\geq 0})$ is again a strictly positive function. Sufficient conditions on r will be found for the solution X of (5.3) to converge almost surely to the (unique) solution $u_0 \in W_0^{1,2}(D)$ of (5.2) (see Example 5.3.1 below).

A common approach to equations like (5.3) is to interpret them in the mild sense, as an equation

$$X_t = U(t,0)y + \int_0^t U(t,s)r(X_s)\,\mathrm{d}s + \int_0^t \alpha(s)U(t,s)\sigma(s,X_s)\,\mathrm{d}W_s,$$

where U is the evolution operator generated by $\alpha(\cdot)\Delta$. However, our proofs rely heavily on the use of Lyapunov functions, while mild solutions are not semimartingales and the Itô formula cannot be applied directly to them. Approximations of a rather technical nature is needed, thence one may very well argue that our use of technically more demanding notion of variational solutions makes things simpler. Moreover, this choice makes it possible to deal with quasilinear problems (see Examples 5.3.2, 5.3.3), which are not amenable to mild solution approach.

Before stating our main results we have to discuss Itô's formula for variational solutions we shall need. This is done in the next sections; the main results are stated and proved in Section 5.2, in Section 5.3 some illustrative examples are provided.

5.1 Itô's formula for variational solutions

Since the process X solving (4.1) is in general only B^*-valued, the Itô formula cannot be used to compute $\varphi(X)$ for an arbitrary $\varphi \in C^2(H)$ and extra assumptions on φ are needed. We state here two Itô formula-type results which we shall need later.

First, let $(\Omega, \mathscr{F}, (\mathscr{F}_t), \boldsymbol{P})$ be a stochastic basis satisfying the usual conditions and carrying a standard cylindrical $(\mathscr{F}_t)$-Wiener process W on K. Assume that:

1° $u_0 \colon \Omega \longrightarrow H$ is an $\mathscr{F}_0$-measurable random variable,

2° $Z \colon \mathbb{R}_{\geq 0} \times \Omega \longrightarrow \mathscr{L}_2(K, H)$ is a progressively measurable process such that $\|Z\|_{\mathscr{L}_2} \in L^2_{\mathrm{loc}}(\mathbb{R}_{\geq 0})$ $\boldsymbol{P}$-almost surely,

3° $v \colon \mathbb{R}_{\geq 0} \times \Omega \longrightarrow B^*$ is a progressively measurable process with $\|v\|_{B^*} \in L^{p/(p-1)}_{\mathrm{loc}}(\mathbb{R}_{\geq 0})$ $\boldsymbol{P}$-almost surely for some $p \in (1, \infty)$,

4° if u is the B^*-valued proces defined by

$$u(t) = u_0 + \int_0^t v(s)\,\mathrm{d}s + \int_0^t Z(s)\,\mathrm{d}W_s, \quad t \geq 0, \tag{5.4}$$

then there exists a B-valued process $\tilde{u}$ such that $\tilde{u} \in L^p_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; B)$ $\boldsymbol{P}$-almost surely and $u = \tilde{u}$ $\lambda \otimes \boldsymbol{P}$-almost everywhere on $\mathbb{R}_{\geq 0} \times \Omega$.

Then u has sample paths in $C(\mathbb{R}_{\geq 0}; H)$ $\boldsymbol{P}$-almost surely and

$$\begin{aligned} \|u(t)\|_H^2 = \|u_0\|_H^2 + \int_0^t \Big\{ 2\langle v(s), \tilde{u}(s)\rangle_{B^*,B} + \|Z_s\|^2_{\mathscr{L}_2} \Big\}\,\mathrm{d}s \\ + 2\int_0^t \langle Z(s)^* u(s), \cdot\rangle_K\,\mathrm{d}W_s \end{aligned}$$

for all $t \geq 0$ $\boldsymbol{P}$-almost surely, see[65] Théorème 3.1 on p. 57,[45] Theorem 2.17 or[54] Theorem 4.2.5.

Comparing this result with Definition 4.1.1 we see that any solution X of (4.1) has path continuous in H $\boldsymbol{P}$-almost surely.

In order to establish the Itô formula for functions more general than $\|\cdot\|_H^2$ one needs an additional hypothesis (A.1.5)

(C) *Both B and B^* are uniformly convex.*

Let $\mathfrak{I}$ be the set of all functions $\varphi \in C^1(H)$ such that the second Gâteaux derivative $D^2\varphi(x)$ exists at all points $x \in H$, the functions φ, $D\varphi$ and $D^2\varphi$

are bounded on bounded sets in H, the mapping $x \longmapsto D^2\varphi(x)$ is continuous from H to $(\mathscr{L}(H), \text{weak}^*)$, the restriction $D\varphi_{|B}$ maps continuously $(B, \|\cdot\|)$ to (B, weak) and there exists a constant $k < \infty$ such that $\|D\varphi(x)\|_B \leq k(1 + \|x\|_B)$ for every $x \in B$. If the process u defined by (5.4) satisfies the hypotheses 1°–4° above and $\varphi \in \mathfrak{I}$ then

$$\varphi(u(t)) = \varphi(u_0)$$

$$+ \int_0^t \Big\{ \big\langle v(s), D\varphi(\tilde{u}(s)) \big\rangle_{B^*,B} + \frac{1}{2} \operatorname{Tr}\big(D^2\varphi(u(s)) Z(s) Z^*(s) \big) \Big\} \,\mathrm{d}s$$

$$+ \int_0^t \big\langle Z(s)^* D\varphi(u(s)), \cdot \big\rangle_K \,\mathrm{d}W_s$$

for all $t \geq 0$ $\boldsymbol{P}$-almost surely, see[65] Théorème 4.2 on p. 65, cf. also[44] Theorem 3.1. In particular, $\varphi(u)$ is a continuous real-valued semimartingale, hence the process $\psi(t, \varphi(u(t)))$ may be expressed by means of the real-valued case of the Itô formula, provided ψ belongs to the set $C^{1,2}$ of all functions $\zeta \in C^1(\mathbb{R}_{\geq 0} \times \mathbb{R})$ such that $\zeta(t, \cdot) \in C^2(\mathbb{R})$ for all $t \geq 0$ and $(t, x) \longmapsto \frac{\partial^2 \zeta}{\partial x^2}(t, x)$ is a continuous function on $\mathbb{R}_{\geq 0} \times \mathbb{R}$. We shall denote by $\mathfrak{K}$ the set of all functions ξ on $\mathbb{R}_{\geq 0} \times H$ of the form $\xi(t, x) = \psi(t, \varphi(x))$, $\psi \in C^{1,2}$, $\varphi \in \mathfrak{I}$. For $\xi \in \mathfrak{K}$ one gets the expected equality

$$\begin{aligned} \xi(t, u(t)) = \xi(0, u_0) + \int_0^t \Big\{ &\frac{\partial \xi}{\partial t}(s, u(s)) + \big\langle v(s), D_x \xi(\tilde{u}(s)) \big\rangle_{B^*,B} \\ &+ \frac{1}{2} \operatorname{Tr}\big(D_x^2 \xi(u(s)) Z(s) Z^*(s) \big) \Big\} \,\mathrm{d}s \\ &+ \int_0^t \big\langle Z(s)^* D_x \xi(u(s)), \cdot \big\rangle_K \,\mathrm{d}W_s. \end{aligned}$$

Note that $D_x \xi(t, x) = \frac{\partial \psi}{\partial x}(t, \varphi(x)) D\varphi(x)$, so the term $\langle v(t), D_x \psi(t, \tilde{u}(t)) \rangle_{B^*,B}$ remains well defined. In the examples below a special case, following from the product rule for semimartingales, is sufficient:

$$\mathrm{d}\big(g(t) \varphi(u(t)) \big) = g'(t) \varphi(u(t)) \,\mathrm{d}t + g(t) \,\mathrm{d}\varphi(u(t))$$

whenever $g \in C^1(\mathbb{R}_{\geq 0})$.

Remark 5.1.1. *a) The hypothesis (C) is obviously satisfied if B is a Hilbert space. Let us emphasize that (C) can be omitted if $\varphi = \|\cdot\|_H^2$ or, more generally, if processes of the form $\psi(t, \|u(t)\|_H^2)$ with $\psi \in C^{1,2}$ are considered.*

b) An Itô formula for the process $\chi(t, u(t))$, where χ is a suitable smooth function on $\mathbb{R}_{\geq 0} \times H$, is proved in[19] *Theorem 7.2.1, but under rather restrictive additional assumptions on u.*

5.2 Main results

Following[61], we derive the convergence of a Robbins-Monro type procedure as an immediate corollary to a theorem providing sufficient conditions for the convergence of path of any solution of (4.1) to a singleton, which will be established first. Hence, let us consider the equation (4.1), that is

$$\mathrm{d}X = f(t, X)\,\mathrm{d}t + \sigma(t, X)\,\mathrm{d}W, \quad X_0 \sim \mu,$$

and denote by L the Kolmogorov operator associated with it, namely, if $h \in \mathfrak{K}$, then we set

$$Lh(t, x) = \frac{\partial h}{\partial t}(t, x) + \big\langle f(t, x), D_x h(t, x)\big\rangle_{B^*, B} + \frac{1}{2}\operatorname{Tr}\big(D_x^2 h(t, x)(\sigma\sigma^*)(t, x)\big),$$

$t \in \mathbb{R}_{\geq 0}$, $x \in B$.

Further, let us consider the following conditions:

(H1) $\varphi\colon \mathbb{R}_{\geq 0} \times H \longrightarrow \mathbb{R}_{\geq 0}$ *is a Borel function and* $x_0 \in H$ *a point such that*

$$\inf_{t \geq 0} \inf_{\|x - x_0\|_H \geq \varepsilon} \varphi(t, x) > 0 \quad \textit{for any } \varepsilon > 0. \tag{5.5}$$

(H2) $V \in \mathfrak{K}$ *is a function satisfying*

$$\lim_{x \to x_0} \sup_{t \geq 0} V(t, x) = 0 \quad \textit{in } H, \tag{5.6}$$

$$\inf_{t \geq 0} \inf_{\|x - x_0\|_H \geq \varepsilon} V(t, x) > 0 \quad \textit{for any } \varepsilon > 0, \tag{5.7}$$

x_0 being the point introduced in (H1)*, and*

$$\int_H V(0,y)\,\mathrm{d}\mu(y) < \infty. \tag{5.8}$$

(H3) $\alpha, \gamma\colon \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{>0}$ *are Borel functions such that* $\alpha \in L^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0}) \setminus L^1(\mathbb{R}_{\geq 0})$*,* $\gamma \in L^1(\mathbb{R}_{\geq 0})$*.*

Now we are prepared to state and prove the main theorem.

Theorem 5.2.1. *Suppose that* (A1) *and* (C) *are satisfied and there exist functions φ, V, α and γ satisfying* (H1)–(H3) *and*

$$LV(t,x) \leq -\alpha(t)\varphi(t,x) + \gamma(t)\big[1 + V(t,x)\big] \quad \textit{for all } t \geq 0,\ x \in B. \tag{5.9}$$

Let $((\Omega, \mathscr{F}, (\mathscr{F}_t), \boldsymbol{P}), W, X)$ *be any solution to* (4.1)*, then*

$$\lim_{t\to\infty} \|X_t - x_0\|_H = 0 \quad \boldsymbol{P}\textit{-almost surely.} \tag{5.10}$$

Remark 5.2.1. *Tracing the proof below one may check easily that – as in the finite-dimensional case – it suffices to assume instead of* (5.5)–(5.7) *that* $V \geq 0$*, there exists* $\tau \geq 0$ *such that*

$$\lim_{x\to x_0} \sup_{t\geq\tau} V(t,x) = 0,$$

and that for any $\varepsilon > 0$ *there exists* $\tau' = \tau'(\varepsilon)$ *such that*

$$\inf_{t\geq\tau'} \inf_{\|x-x_0\|_H\geq\varepsilon} V(t,x) > 0, \quad \inf_{t\geq\tau'} \inf_{\|x-x_0\|_H\geq\varepsilon} \varphi(t,x) > 0.$$

Remark 5.2.2. *The singleton* $\{x_0\}$ *may be replaced with an arbitrary closed set* $\Gamma \subseteq H$*. Let* (5.5)–(5.7) *be modified in the following way:*

$$\lim_{\mathrm{dist}(x,\Gamma)\to 0} \sup_{t\geq 0} V(t,x) = 0 \quad \textit{in } H,$$
$$\inf_{t\geq 0} \inf_{\mathrm{dist}(x,\Gamma)\geq\varepsilon} V(t,x) \wedge \varphi(t,x) > 0 \quad \textit{for all } \varepsilon > 0$$

and let $V = 0$ *on* $\mathbb{R}_{\geq 0} \times \Gamma$*. Then* $\mathrm{dist}(X_t, \Gamma) \longrightarrow 0$ *as* $t \to \infty$ $\boldsymbol{P}$*-almost* *surely. The proof requires only very straightforward changes; unfortunately, this result is usually too weak to be applied to equations with multiple roots (cf. the discussion in*[61]*, Chapter 5).*

Proof. a) The first two steps of the proof are essentially known from stability theory of stochastic PDEs, but we provide them for completeness and as we shall refer to parts of the argument in the sequel. Set

$$U(t,x) = \exp\Big(\int_t^\infty \gamma(r)\,\mathrm{d}r\Big)\big[1+V(t,x)\big], \quad (t,x) \in \mathbb{R}_{\geq 0} \times H.$$

Since $\gamma \in L^1(\mathbb{R}_{\geq 0})$, U is obviously well defined, $U \geq 0$ on $\mathbb{R}_{\geq 0} \times H$ and $U \in \mathfrak{K}$. An easy calculation shows that

$$LU(t,x) \leq -\alpha(t)\varphi(t,x) \quad \text{for all } t \geq 0,\ x \in B, \tag{5.11}$$

in particular $LU \leq 0$ on $\mathbb{R}_{\geq 0} \times B$.

b) We aim at proving that $(U(t,X_t),\ t \geq 0)$ is a supermartingale. To this end, set

$$\tau_n = \inf\big\{t \geq 0;\ \|X_t\|_H \geq n\big\} \wedge \inf\Big\{t \geq 0;\ \int_0^t \|\sigma(r,\tilde{X}_r)\|^2_{\mathscr{L}_2} \geq n\Big\},\ n \in \mathbb{N}$$

(with the convention $\inf \emptyset = +\infty$), where $\tilde{X}$ is the process introduced in Definition 4.1.1. Plainly, $\tau_n \nearrow \infty$ as $n \to \infty$ $\boldsymbol{P}$-almost surely. Using the Itô formula and (5.11) we get

$$\begin{aligned}
&U(t\wedge\tau_n, X(t\wedge\tau_n)) - U(0,X_0)\\
&= \int_0^{t\wedge\tau_n} LU(r,\tilde{X}_r)\,\mathrm{d}r + \int_0^{t\wedge\tau_n} \langle \sigma(r,\tilde{X}_r)^* D_x U(r,X_r), \cdot\rangle_K \,\mathrm{d}W_r\\
&\leq \int_0^{t\wedge\tau_n} \langle \sigma(r,\tilde{X}_r)^* D_x U(r,X_r), \cdot\rangle_K \,\mathrm{d}W_r.
\end{aligned}$$

Note that

$$\boldsymbol{E}\int_0^{t\wedge\tau_n} \|\sigma(r,\tilde{X}_r)^* D_x U(r,X_r)\|_K^2\,\mathrm{d}r < \infty$$

for any $t \geq 0$ due to the definition of τ_n and boundedness of $D_x U$ on bounded subsets of $\mathbb{R}_{\geq 0} \times H$, as

$$\begin{aligned}&\int_0^{t\wedge\tau_n} \|\sigma(r,\tilde{X}_r)^* D_x U(r,X_r)\|_K^2 \,\mathrm{d}r \\ &\leq \sup_{\substack{0\leq r\leq t\\ \|z\|_H\leq n}} \|D_x U(r,z)\|_H^2 \int_0^{t\wedge\tau_n} \|\sigma(r,\tilde{X}_r)\|_{\mathscr{L}_2}^2 \,\mathrm{d}r.\end{aligned}$$

We see that

$$\int_0^{\cdot\wedge\tau_n} \langle \sigma(r,\tilde{X}_r)^* D_x U(r,X_r), \cdot\rangle_K \,\mathrm{d}W_r$$

is a martingale, hence

$$\boldsymbol{E} U(t\wedge\tau_n, X(t\wedge\tau_n)) \leq \boldsymbol{E} U(0,X_0) \leq e^{\|\gamma\|_{L^1}} \boldsymbol{E} V(0,X_0) < \infty$$

for all $t \geq 0$ and $n \in \mathbb{N}$ by (5.8). Since $U \in C(\mathbb{R}_{\geq 0} \times H)$ and the paths of X are continuous in H, we obtain

$$\begin{aligned}\boldsymbol{E} U(t,X_t) &= \boldsymbol{E} \lim_{n\to\infty} U(t\wedge\tau_n, X(t\wedge\tau_n)) \leq \liminf_{n\to\infty} \boldsymbol{E} U(t\wedge\tau_n, X(t\wedge\tau_n)) \\ &\leq \boldsymbol{E} U(0,X_0)\end{aligned}$$

by the Fatou lemma. Thus $U(t,X_t) \in L^1(\boldsymbol{P})$ for every $t \in \mathbb{R}_{\geq 0}$. Analogously, for any $0 \leq s \leq t$, we have

$$U(t\wedge\tau_n, X(t\wedge\tau_n)) - U(s\wedge\tau_n, X(s\wedge\tau_n)) \leq \int_{s\wedge\tau_n}^{t\wedge\tau_n} \langle \sigma(r,\tilde{X}_r)^* D_x U(r,X_r), \cdot\rangle_K \,\mathrm{d}W_r;$$

so

$$\boldsymbol{E}\big[U(t\wedge\tau_n, X(t\wedge\tau_n)) \,\big|\, \mathscr{F}_s\big] - U(s\wedge\tau_n, X(s\wedge\tau_n) \leq 0.$$

The Fatou lemma for conditional expectations now implies that

$$\begin{aligned}\boldsymbol{E}\big[U(t,X_t) \,\big|\, \mathscr{F}_s\big] &\leq \liminf_{n\to\infty} \boldsymbol{E}\big[U(t\wedge\tau_n, X(t\wedge\tau_n)) \,\big|\, \mathscr{F}_s\big] \\ &\leq \liminf_{n\to\infty} U(s\wedge\tau_n, X(s\wedge\tau_n)) \\ &= U(s,X_s)\end{aligned}$$

$\boldsymbol{P}$-almost surely, which is the supermartingale property. For further use, let us note that proceeding as above, we get

$$
\begin{aligned}
-\boldsymbol{E}\int_0^{t\wedge\tau_n} LU(r,\tilde{X}_r)\,\mathrm{d}r &= \boldsymbol{E}U(0,X_0)-\boldsymbol{E}U(t\wedge\tau_n, X(t\wedge\tau_n)) \\
&\le \boldsymbol{E}U(0,X_0),
\end{aligned}
$$

whence

$$
-\boldsymbol{E}\int_0^{\infty} LU(r,\tilde{X}_r)\,\mathrm{d}r \le \boldsymbol{E}U(0,X_0) < \infty, \tag{5.12}
$$

again by the Fatou lemma.

Since $U(t,X_t)$ is a continuous nonnegative supermartingale, the martingale convergence theorem yields a random variable $U_\infty \in L^1(\boldsymbol{P})$ such that

$$
\lim_{t\to\infty} U(t,X_t) = U_\infty \quad \boldsymbol{P}\text{-almost surely.}
$$

From the definition of U it follows that there exists $\Omega_s \in \mathscr{F}$, $\boldsymbol{P}(\Omega_s)=1$, such that $1 \le U_\infty(\omega) < \infty$ and

$$
\lim_{t\to\infty} V(t, X(t,\omega)) = V_\infty(\omega) \equiv U_\infty(\omega) - 1
$$

for any $\omega \in \Omega_s$.

c) Since

$$
\int_0^\infty \alpha(r)\varphi(r,\tilde{X}_r)\,\mathrm{d}r \le -\int_0^\infty LU(r,\tilde{X}_r)\,\mathrm{d}r \quad \text{on } \Omega
$$

by (5.11), the integral on the right-hand side is a nonnegative random variable with a finite expectation by (5.12), and $X=\tilde{X}$ λ-almost everywhere on $\mathbb{R}_{\ge 0}$ $\boldsymbol{P}$-almost surely, there exists $\Omega_i \in \mathscr{F}$, $\boldsymbol{P}(\Omega_i)=1$, such that

$$
\int_0^\infty \alpha(r)\varphi(r, X(r,\omega))\,\mathrm{d}r < \infty
$$

for every $\omega \in \Omega_i$.

d) Now we check that

$$
\liminf_{t\to\infty} \|X(t,\omega) - x_0\|_H = 0 \tag{5.13}
$$

for all $\omega \in \Omega_i$. Striving for a contradiction assume that $\omega \in \Omega_i$ but

$$\liminf_{t\to\infty} \|X(t,\omega) - x_0\|_H > 0.$$

Then there exists $t_0 \in \mathbb{R}_{\geq 0}$ and $\varepsilon > 0$ such that $\|X(t,\omega) - x_0\|_H \geq \varepsilon$ for any $t \geq t_0$; by (H1) we may find $\delta > 0$ satistying $\varphi(t, X(t,\omega)) \geq \delta$ for all $t \geq t_0$, therefore

$$\int_{t_0}^{\infty} \alpha(r)\varphi(r, X(r,\omega))\,\mathrm{d}r \geq \delta \int_{t_0}^{\infty} \alpha(r)\,\mathrm{d}r = +\infty$$

by (H3), however, this contradicts the definition of Ω_i.

e) It remains to prove that

$$\lim_{t\to\infty} \|X(t,\omega) - x_0\|_H = 0 \quad \text{for all } \omega \in \Omega_i \cap \Omega_s. \tag{5.14}$$

Assume that $\omega \in \Omega_i \cap \Omega_s$ but (5.14) fails. Then there exist $t_n \nearrow \infty$ and $\varepsilon > 0$ such that $\|X(t_n,\omega) - x_0\|_H \geq \varepsilon$. By (5.7), a $\eta > 0$ may be found such that $V(t_n, X(t_n,\omega)) \geq \eta$, consequently

$$\eta \leq \lim_{n\to\infty} V(t_n, X(t_n,\omega)) = V_\infty(\omega).$$

On the other hand, by (5.13) there exist $s_n \nearrow \infty$ such that

$$\|X(s_n,\omega) - x_0\|_H \to 0 \text{ as } n \to \infty,$$

hence

$$0 \leq V_\infty(\omega) = \lim_{n\to\infty} V(s_n, X(s_n,\omega)) \leq \lim_{n\to\infty} \sup_{r\geq 0} V(r, X(s_n,\omega)) = 0$$

by (5.6). This contradiction proves (5.14) and the proof of Theorem 5.2.1 is completed. Q.E.D.

Remark 5.2.3. *By* (5.6) *and Theorem 5.2.1,*

$$\lim_{t\to\infty} V(t, X_t) = 0 \quad \boldsymbol{P}\text{-almost surely.} \tag{5.15}$$

The estimate

$$\boldsymbol{E}V(t,X_t) = e^{-\int_t^\infty \gamma\,\mathrm{d}r}\boldsymbol{E}U(t,X_t) - 1 \le \boldsymbol{E}U(0,X_0) < \infty, \quad t \ge 0,$$

was established in the course of the proof. Therefore, if $\nu \in (0,1)$ *then the set* $\{V(t,X_t)^\nu,\, t \ge 0\}$ *is uniformly integrable and* (5.15) *implies that*

$$\lim_{t\to\infty} \boldsymbol{E}V(t,X_t)^\nu = 0.$$

Now we may proceed to a theorem on stochastic approximation.

Corollary 5.2.2. *Let* (C) *be satisfied, let* $R\colon B \longrightarrow B^*$ *and* $\sigma\colon \mathbb{R}_{\ge 0} \times B \longrightarrow \mathscr{L}_2(K,H)$ *be Borel function and* μ *a Borel probability measure on* H*. Let* $x_0 \in B$ *be such that* $R(x_0) = 0$*. Suppose that there exist* $V \in \mathfrak{I} \cap L^1(\mu)$ *and a Borel function* $\varphi\colon H \longrightarrow \mathbb{R}_{\ge 0}$ *satisfying*

$$V(x_0) = 0, \qquad \inf_{\|x-x_0\|_H \ge \varepsilon} \bigl\{V(x) \wedge \varphi(x)\bigr\} > 0 \quad \textit{for any } \varepsilon > 0,$$

$$\bigl\langle R(x), DV(x)\bigr\rangle_{B^*,B} \le -\varphi(x) \quad \textit{for all } x \in B, \tag{5.16}$$

and

$$\operatorname{Tr}\bigl(D^2V(x)(\sigma\sigma^*)(t,x)\bigr) \le K\bigl(1 + V(x)\bigr) \tag{5.17}$$

for some $K < \infty$ *and all* $(t,x) \in \mathbb{R}_{\ge 0} \times B$*.*

Let $\alpha\colon \mathbb{R}_{\ge 0} \longrightarrow \mathbb{R}_{>0}$ *be a Borel function such that*

$$\int_0^\infty \alpha(r)\,\mathrm{d}r = \infty, \qquad \int_0^\infty \alpha^2(r)\,\mathrm{d}r < \infty.$$

Then any solution $(\Omega, \mathscr{F}, (\mathscr{F}_t), \boldsymbol{P}), W, X)$ *of the equation*

$$\mathrm{d}X(t) = \alpha(t)R(X(t))\,\mathrm{d}t + \alpha(t)\sigma(t,X(t))\,\mathrm{d}W(t), \quad X(0) \sim \mu, \tag{5.18}$$

satisfies

$$\lim_{t\to\infty} \|X(t) - x_0\|_H = 0 \quad \boldsymbol{P}\textit{-almost surely.}$$

If, moreover, $V(x) \geq L\|x - x_0\|_H^2$ *for some* $L \in \mathbb{R}_{>0}$ *and all* $x \in H$*, then*

$$\lim_{t\to\infty} \boldsymbol{E}\|X(t) - x_0\|^\nu = 0$$

for any $\nu \in (0, 2)$*.*

Remark 5.2.4. *a) Note that* (5.16) *may be satisfied only if* x_0 *is the unique root of* R*.*

b) As in Theorem 5.2.1, we do not assume that there exists a unique solution of (5.18)*, we only claim that if a solution exists, then it converges to the root of* R*. Of course, in examples the problem of existence and uniqueness of solutions gains prominence.*

5.3 Examples

Example 5.3.1. Let $\tilde{D} \subseteq \mathbb{R}^d$ be a bounded open set with a sufficiently smooth boundary $\partial\tilde{D}$, $g\colon \mathbb{R} \longrightarrow \mathbb{R}$ a Borel function and f a (generalized) function on $\tilde{D}$. Let us consider a nonlinear elliptic equation

$$\Delta u + g(u) = f \text{ in } \tilde{D}, \quad u = 0 \text{ on } \partial\tilde{D}. \tag{5.19}$$

Set $H = L^2(\tilde{D})$, $B = W_0^{1,2}(\tilde{D})$ and denote by G the superposition operator defined by g. Assume that G is a continuous mapping from B to H and that there exists $\varrho \in \mathbb{R}$ such that

$$\begin{aligned} &\big\langle G(u) - G(v), u - v \big\rangle_H \leq \varrho\|u - v\|_H^2, \\ &\big\langle G(u), u \big\rangle_H \leq \varrho\big(1 + \|u\|_H^2\big), \quad \|G(u)\|_H \leq \varrho\big(1 + \|u\|_B\big) \end{aligned} \tag{5.20}$$

for all $u, v \in B$. Note that (5.20) is surely satisfied if g is either Lipschitz continuous or nonincreasing. Let $\sigma\colon \mathbb{R}_{\geq 0} \times B \longrightarrow \mathscr{L}_2(K, H)$ be a Borel function such that

$$\sup_{0\leq t\leq T} \sup_{x\in B} \frac{\|\sigma(t,x)\|_{\mathscr{L}_2}}{1 + \|x\|_H} + \sup_{0\leq t\leq T} \sup_{\substack{x,y\in B \\ x\neq y}} \frac{\|\sigma(t,x) - \sigma(t,y)\|_{\mathscr{L}_2}}{\|x - y\|_H} < \infty \tag{5.21}$$

for any $T \in \mathbb{R}_{\geq 0}$. Finally, let $\alpha \colon \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{>0}$ satisfy

$$\alpha \in L^2(\mathbb{R}_{\geq 0}) \setminus L^1(\mathbb{R}_{\geq 0}) \quad \text{and} \quad 0 < \inf_{[0,T]} \alpha \leq \sup_{[0,T]} \alpha < \infty \tag{5.22}$$

for any $T \in \mathbb{R}_{\geq 0}$, let $f \in B^*$ and let μ be a Borel probability measure on H with a finite second moment, i.e. $\|\cdot\|_H \in L^2(\mu)$. Then it may be checked easily that all hypotheses of Theorem 4.2.1 (as in Example 4.2.1) are satisfied and hence there exists a unique solution $((\Omega, \mathscr{F}, (\mathscr{F}_t), \boldsymbol{P}), W, X)$ to the stochastic parabolic equation

$$\mathrm{d}X = \alpha(t)\big(\Delta X + G(X) - f\big)\,\mathrm{d}t + \alpha(t)\sigma(t, X)\,\mathrm{d}W(t), \quad X(0) \sim \mu,$$

the Dirichlet Laplacian Δ being interpreted as an operator in $\mathscr{L}(B, B^*)$ in a natural way. Assume that there exists a weak solution $u_0 \in B$ of (5.19); one may consult e.g.[3],[67] Chapter 9 or references therein for results in this direction. We want to apply Corollary 5.2.2 with $V = \|\cdot - u_0\|_H^2$. Since u_0 solves (5.19),

$$\big\langle \Delta u + G(u) - f, DV(u)\big\rangle_{B^*,B} = 2\big\langle \Delta(u - u_0) + G(u) - G(u_0), u - u_0\big\rangle_{B^*,B}$$

and it is known that

$$\big\langle \Delta(u - u_0), u - u_0\big\rangle_{B^*,B} \leq -\varkappa\|u - u_0\|_H^2$$

for some $\varkappa > 0$ and all $u \in B$, so Corollary 5.2.2 implies that

$$\lim_{t\to\infty} \|X(t) - u_0\|_H = 0 \ \boldsymbol{P}\text{-almost surely} \quad \text{and} \quad \lim_{t\to\infty} \boldsymbol{E}\|X(t) - u_0\|^{2-\varepsilon} = 0 \tag{5.23}$$

for all $\varepsilon \in (0, 2)$, provided

$$\big\langle G(u) - G(u_0), u - u_0\big\rangle_H \leq (\varkappa - \eta)\|u - u_0\|_H^2 \tag{5.24}$$

for some $\eta > 0$ and all $u \in B$, and

$$\sup_{t\geq 0} \sup_{x\in B} \frac{\|\sigma(t, x)\|_{\mathscr{L}_2}}{1 + \|x\|_H} < \infty. \tag{5.25}$$

As we have already mentioned, (5.24) is satisfied if g is either nonincreasing, or Lipschitz continuous with a sufficiently small Lipschitz constant.

Example 5.3.2. Let $\tilde{D} \subseteq \mathbb{R}^d$ be a bounded domain with a sufficiently smooth boundary and $p \in (2, \infty)$. Set $B = W_0^{1,p}(\tilde{D})$ and $H = L^2(\tilde{D})$, we shall consider the p-Laplacian

$$\Delta_p u = \operatorname{div}\bigl(|\nabla u|^{p-2} \nabla u\bigr),$$

that is, rigorously, an operator $\Delta_p \colon B \longrightarrow B^*$ defined by

$$\langle \Delta_p u, v \rangle_{B^*,B} = -\int_D |\nabla u(r)|^{p-2} \langle \nabla u(r), \nabla v(r) \rangle \,\mathrm{d}r, \quad u, v \in B.$$

Let $f \in H$. It follows from[3] Theorem 2.6.8 that the quasilinear elliptic equation

$$\Delta_p u = f \text{ in } \tilde{D}, \quad u = 0 \text{ on } \partial\tilde{D} \tag{5.26}$$

has a unique weak solution $u_0 \in B$. Likewise, the stochastic equation

$$\mathrm{d}X = \alpha(t)\bigl(\Delta_p X - f\bigr)\,\mathrm{d}t + \alpha(t)\sigma(t, X)\,\mathrm{d}W(t), \quad X(0) \sim \mu$$

has a unique variational solution $((\Omega, \mathscr{F}, (\mathscr{F}_t), \boldsymbol{P}), W, X)$ if α and σ satisfy (5.22) and (5.21), respectively, and μ is a Borel probability measure on H with a finite second moment, see Example 4.2.2. Again we shall use Corollary 5.2.2 with $V = \|\cdot - u_0\|_H^2$. Due to the inequality

$$\bigl\langle \|t\|^{p-2} t - \|s\|^{p-2} s, t - s \bigr\rangle \geq c_p \|t - s\|^p \quad \text{for a } c_p > 0 \text{ and all } s, t \in \mathbb{R}^d \tag{5.27}$$

(see e.g.[75] p. 210) the operator $-\Delta_p$ is strongly monotone,

$$\langle \Delta_p u - \Delta_p v, u - v \rangle_{B^*,B} =$$
$$-\int_{\tilde{D}} \bigl\langle \|\nabla u(r)\|^{p-2} \nabla u(r) - \|\nabla v(r)\|^{p-2} \nabla v(r), u(r) - v(r) \bigr\rangle \,\mathrm{d}r \leq -c_p \|u - v\|_B^p$$

for all $u, v \in B$, whence we have

$$\begin{aligned}\langle \Delta_p u - f, DV(u)\rangle_{B^*,B} &= 2\langle \Delta_p u - \Delta_p u_0, u - u_0\rangle_{B^*,B} \\ &\leq -2c_p \|u - u_0\|_B^p \\ &\leq -\tilde{c}\,\|u - u_0\|_H^p\end{aligned}$$

for some constant $\tilde{c} > 0$ and all $u \in B$. Therefore, if σ satisfies (5.25) then (5.23) holds true for all $\varepsilon \in (0, 2)$.

Example 5.3.3. In this example equations involving the porous medium operator (Example 4.2.3) will be considered. Let $\tilde{D} \subseteq \mathbb{R}^d$ be a bounded domain with a sufficiently smooth boundary and $p \in (2, \infty)$, set $B = L^p(\tilde{D})$, $H = (W_0^{1,2}(\tilde{D}))^*$ and $\Psi(s) = s|s|^{p-2}$ for $s \in \mathbb{R}$, and define

$$A\colon B \longrightarrow B^*,\ u \longmapsto \Delta\Psi(u).$$

Let $\sigma\colon \mathbb{R}_{\geq 0} \times B \longrightarrow \mathscr{L}_2(K, H)$ and $\alpha\colon \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{>0}$ satisfy (5.21) and (5.22), respectively, let $f \in B^*$ and let μ be a Borel probability measure on H with a finite second moment. Then there exists a unique solution X of

$$\mathrm{d}X = \alpha(t)\big(A(X) - f\big)\,\mathrm{d}t + \alpha(t)\sigma(t, X)\,\mathrm{d}W(t), \quad X(0) \sim \mu.$$

Using the inequality (5.27) with $d = 1$ one may check that $-A$ is strongly monotone:

$$\begin{aligned}\langle A(u) - A(v), u - v\rangle_{B^*,B} &= \langle \Delta(\Psi(u) - \Psi(v)), u - v\rangle_{B^*,B} \\ &= -\int_{\tilde{D}} (\Psi(u(r)) - \Psi(v(r)))(u(r) - v(r))\,\mathrm{d}r \\ &\leq -c_p\|u - v\|_B^p.\end{aligned}$$

It follows, first, that the problem $Au = f$ has a unique solution $u_0 \in B$ and, secondly, choosing $V = \|\cdot - u_0\|_H^2$ we get

$$\langle A(u) - f, DV(u)\rangle_{B^*,B} \leq -\hat{c}\,\|u - u_0\|_H^p$$

for some $\hat{c} > 0$ and any $u \in B$. Therefore, (5.23) holds provided (5.25) is satisfied.

6

Unsolved questions related to the thesis

There are several direction, where one might improve the results contained in chapters two and three. As the most important or intriguing we mention

- There is the question of uniqueness of solution to martingale problem we considered in Theorem 3.1.7. As mentioned previously, uniqueness of martingale problems in infinite dimension is notoriously hard question with very few satisfying results so far.
- We manage to prove the uniqueness of our approximating procedure, which enables us to construct proper Markov process. More crucial here is the question of ergodicity. That is, whether the invariant measure we constructed in Theorem 3.3.2 is unique. Since our assumptions don't tie the size of interactions with the space where process lives, one is tempted to believe that the answer is no. However, we have no rigorous answer in this direction.
- Possibly just a curiosity, but nevertheless nice exercise in Analysis seems the question about which spaces the condition **(H5)** allow. Ob-

viously, we see that if the weights $w(i)$ decays faster than $\frac{1}{n^{2+\epsilon}}$ for $\epsilon > 0$, the space is large enough, since one can take $v(i)$ to be $\frac{1}{n^{1+\frac{\epsilon}{2}}}$. We would be interested, whether one can say more. Perhaps it would be nice, to at least decide, whether weights $\frac{1}{n^{1+\alpha}}, \alpha \in (0, 1]$ can satisfy this condition or not

- Tracing our proofs, one see that seemingly artificial restrictions to interactions **(H2)** stems from our proof of uniqueness of approximating procedure. The solution to martingale problem itself, we can construct without any such restrictions. Hence the question arises, whether one could devise some more inventive method, how to rigorously proof uniqueness under less restrictive assumptions.

The results obtained in chapter five are more complete and in a way provide satisfying extension of continuous Robbins-Monro procedure to infinite dimensional setting. Nevertheless, one can naturally ask, if some other procedures could find its infinite dimensional analogue. Concretely let us note

- Instead of searching for root of an unknown function, one can try to use stochastic approximation to locate maximum point x_0 of an unknown function f. The method to do so was suggested by Kiefer and Wolfowitz in discrete setting in[39]. Nevel'son and Khas'minskiĭ in[61] pp. 94 describe continuous version of Kiefer-Wolfowitz procedure. To present their approach, we let $f = f(x_1, \cdots, x_d)$ be smooth function, which we cannot directly observe, but have advance knowledge, that its maximum exists and is unique. In another words one attempts to solve equation $\nabla f = 0$. Let $(e_i)_{1\leq i\leq d}$ be basis in $\mathbb{R}^d$ and let $\nabla_c f$ denote vector with coordinates $\frac{f(x+c(t)e_i)-f(x-c(t)e_i)}{2c(t)}$ for some non-zero function $c : [0, \infty) \to \mathbb{R}$. Then Kiefer-Wolfowitz procedure leads to (compare with equation (5.1)) the stochastic differential equation

$$dX = a(t)\nabla_c f(X)dt + \frac{a(t)}{c(t)}\sigma(t, x)dW. \tag{6.1}$$

  Under several conditions on the coefficients it can be proved that indeed $X(t)$ a. s. converges to the maximum x_0. To find some non-trivial infinite dimensional application of these ideas would clearly provide nice extension of our results.

A

Auxiliary theorems and concepts

The order of appearance of auxiliary results follows its appearance in the text.

A.1 Definitions of concepts

Definition A.1.1 (Linear growth and locally Lipschitz condition)**.** Let $b : \mathbb{R}^u \to \mathbb{R}^u$, $\sigma : \mathbb{R}^u \to \mathbb{R}^{d\times u}$, $u, d \in \mathbb{N}$ be Borel measurable functions. We say b and σ satisfy linear growth condition, if

$$\exists K > 0 : \ \|b(x)\|_u \vee \|\sigma\|_{d\times u} \leq K(1 + \|x\|_u) \ \forall x \in \mathbb{R}^u.$$

We say b and σ are locally Lipschitz, provided that for any bounded open $A \subset \mathbb{R}^u$ we have

$$\begin{aligned}\exists K > 0 : \|b(x) - b(y)\|_u &\leq K\|x - y\|_u \ \forall x, y \in A\\ \|b(x) - b(y)\|_{d\times u} &\leq K\|x - y\|_u \ \forall x, y \in A.\end{aligned}$$

b and σ are globally Lipschitz, provided the condition holds for any $x, y \in \mathbb{R}^u$

Definition A.1.2 (Cylindrical Wiener process)**.** Let U be Hilbert space and Q bounded, self-adjoint, strictly positive, i. e. $Qx \neq 0$ for $x \neq 0$, operator on U. Let $U_0 = Q^{1/2}(U)$ with the induced norm $\|u\|_0 = \|Q^{-1/2}(u)\|, u \in U_0$, and let U_1 be an arbitrary Hilbert space such that U is embedded continuously into U_1 and the embedding of U_0 into U_1 is Hilbert-Schmidt. Let $\{e_j\}$ be an orthonormal basis in U_0 and $\{B_j\}$ a family of independent real valued standard Wiener processes. The formula

$$W(t) = \sum_{j=1}^{\infty} e_j B_j(t),\ \ t \geq 0$$

is convergent almost surely in space U_1. It holds that

$$E\langle u, W(t)\rangle\langle v, W(s)\rangle = (t \wedge s)\langle Qu, v\rangle \ u, v \in U.$$

If tr $Q = \infty$ then W is called cylindrical Wiener process, for $Q = I$ we speak about standard cylindrical Wiener process, see[23] pp. 96 for more thorough discussion.

Definition A.1.3 (Hilbert-Schmidt operators)**.** Let $T : K \to H$ be bounded linear operator between Hilbert spaces K and H. T is called Hilbert-Schmidt, if

$$\sum_{k \in \mathbb{N}} \|Te_k\| < \infty,$$

where $e_k, k \in \mathbb{N}$ is an orthonormal basis of K.

Definition A.1.4 (Sobolev space with Dirichlet boundary condition)**.** Let $D \subseteq \mathbb{R}^d$ be bounded open set. For $u \in C_c^\infty(D)$ and $p > 1$ define

$$\|u\|_{1,p} = \left(\int (|u(x)|^p + |\nabla u(x)|^p dx \right)^{1/p}.$$

Then set

$$W_0^{1,p}(D) := \text{completion of } C_c^\infty(D) \text{with respect to } \| \cdot \|_{1,p}.$$

$W_0^{1,p}$ is called the Sobolev space of order one with Dirichlet boundary condition.

Note that there is natural extension of operator $\nabla : C_c^\infty(D) \to L^2(D)$ to domain $W_0^{1,p}(D)$. Indeed, if $u_n, u \in W_0^{1,p}, n \in \mathbb{N}$ with $\|u_n - u\|_{1,p} \to 0$, then $\{\nabla u_n\}$ is a Cauchy sequence in $L^p(D)$, hence there exists limit of this sequence independent of the choice of sequence $\{u_n\}$.

Definition A.1.5 (Uniformly convex space)**.** A uniformly convex space is a normed vector space B such that, for every $0 < \epsilon \leq 2$ there is some $\delta > 0$ such that for any two vectors $u, v \in B$ with $\|u\| = \|v\| = 1$, the condition $\|u - v\| \geq \epsilon$ implies that

$$\left\| \frac{u+v}{2} \right\| \leq 1 - \delta.$$

For our usage it suffices to note that

- Every Hilbert space is uniformly convex
- Every closed subspace of a uniformly convex Banach space is uniformly convex
- $L^p, 1 < p < \infty$ is uniformly convex, as observed first by Clarkson in[20].

A.2 List of theorems used

Theorem A.2.1 (Doob's upcrossing lemma)**.** *Let X_k be a supermartingale and $a < b, a, b \in \mathbb{R}$ given points. Define recursively times*

$$S_1 = \min\{k : X_k \leq a\} \quad T_1 = \min\{k > S_1 : X_k \geq b\}$$

$$S_{i+1} = \min\{k > T_i : X_k \leq a\} \quad T_{i+1} = \min\{k > S_{i+1} : X_k \geq b\}.$$

The number of upcrossings U_n before time n is $U_n = \max\{j : T_j \leq n\}$. Then

$$EU_n \leq \frac{1}{b-a} E[(M_n - a)^-]. \tag{A.1}$$

Proof. See[8] Theorem A.34 □

Theorem A.2.2 (Girsanov theorem)**.** *Let X and Y be solutions to stochastic differential equations*

$$dX(t) = b(X(t))dt + \sigma(X(t))dW_t$$
$$dY(t) = \tilde{b}(Y(t))dt + \sigma(Y(t))dW_t,$$

where the coefficients satisfy condition (A.1.1). Suppose that there exists bounded Borel function $u : \mathbb{R}^d \to \mathbb{R}$, satisfying

$$\sigma u = \tilde{b} - b.$$

Let P_t be the transition function corresponding to X_t, Q_t associated with Y_t. We then have

$$supp\ P_t(x, \cdot) = supp\ Q_t(x, \cdot)\ \forall t > 0, \forall x \in \mathbb{R}^u.$$

Proof. See[64] pp. 166. □

Girsanov theorem is usually stated with much weaker assumptions about u, however this version is sufficient for our purposes.

Theorem A.2.3 (Portmanteau theorem)**.** *Let $\{P_n\}, P$ be probability measures on metric space $\mathcal{S}$. The following conditions are equivalent*

(i) $P_n \xrightarrow{w} P$

(ii) $\limsup_n P_n(F) \leq P(F)\ \forall F$ *closed.*

(iii) $\liminf_n P_n(G) \geq P(G)\ \forall G$ *open.*

Proof. See[11] pp. 16. □

Theorem A.2.4 (Stolz-Cesàro theorem)**.** *Let $\{b_n\}$ be sequence that is strictly increasing and $\lim_{n\to\infty} b_n = \infty$. Let $\{a_n\}$ be a given sequence and assume*

$$\lim_{n\to\infty} \frac{a_{n+1} - a_n}{b_{n+1} - b_n} = l,\ \ l \in \mathbb{R}.$$

Then

$$l = \lim_{n\to\infty} \frac{a_n}{b_n}.$$

Proof. See[60] pp. 85. □

Theorem A.2.5 (Poincaré's inequality)**.** *Suppose that $p \geq 1$ and D is bounded open set in $\mathbb{R}^d$. Then there exists a constant $C > 0$ (depending on D, p and dimension of the problem d) such that*

$$\forall u \in W_0^{1,p}(D) \quad \|u\|_{L^p(D)} \leq C\|\nabla\|_{L^p(D)}.$$

Proof. See[16] pp. 290. □

References

[1] N. Arcozzi, A. Baldi, From Grushin to Heisenberg via an isoperimetric problem. J. Math. Anal. Appl. **340** (2008), no. 1, 165-–174

[2] R. Athreya et al., *Infinite dimensional stochastic differential equations of Ornstein - Uhlenbeck type*, Stochastic Process. Appl. **116** (2006), no. 3, 381–406

[3] M. Badiale, E. Serra, *Semilinear elliptic equations for beginners*, Springer, London 2011

[4] D. Bakry et al., *On gradient bounds for the heat kernel on the Heisenberg group*, Journal of Functional Analysis **255** (2008) 1905–1938

[5] J. Bao, G. Yin, C. Yuan, *Ergodicity for functional stochastic differential equations and applications*, Nonlinear Anal. **98** (2014) 66–82

[6] K. Barty, J.-S. Roy, C. Strugarek, *Hilbert-valued perturbed subgradient algorithms*, Math. Oper. Res. **32** (2007), 551–562

[7] R. Bass, E. Perkins, *On uniqueness in law for parabolic SPDEs and infinite-dimensional SDEs*, Electron. J. Probab. **17** (2012), no. 36, 54 pp.

[8] R. Bass, *Stochastic Processes*, Cambridge University Press, 2011

[9] A. Benveniste, M. Métivier, P. Priouret, *Adaptive algorithms and stochastic approximations*, Springer, Berlin 1990

[10] S. Bhatnagar, H. L. Prasad, L. A. Prashanth, *Stochastic recursive algorithms for optimization*, Lecture Notes in Control Inform. Sci. 434, Springer, London 2013

[11] P. Billingsley, *Convergence of Probability Measures*, 2nd. edition, John Wiley & Sons, 1999

[12] R. Blumenthal, R. Getoor, *Markov processes and potential theory*, Pure and Applied Mathematics, Vol. 29 Academic Press, New York-London 1968

[13] N. N. Bogoliubov, N. M. Krylov, *La theorie generalie de la mesure dans son application a l'etude de systemes dynamiques de la mecanique non-lineaire*, Ann. Math. II (in French) (Annals of Mathematics) **38** (1) (1937), 65—113

[14] A. Bonfiglioli, E. Lanconelli, F. Uguzonni, *Stratified Lie Groups and Potential Theory for Their Sub-Laplacians*, Springer-Verlag Berlin Heidelberg, 2007

[15] V. S. Borkar, *Stochastic approximation: A dynamical systems viewpoint*, Cambridge University Press, Cambridge 2008

[16] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2011

[17] H.-F. Chen, *Continuous-time stochastic approximation: Convergence and asymptotic efficiency*, Stochastics Stochastic Rep. **51** (1994), 217–239

[18] H.-F. Chen, *Stochastic approximation and its applications*, Kluwer, Dordrecht 2002

[19] P.-L. Chow, *Stochastic partial differential equations*, 2nd ed., CRC Press, Boca Raton 2015

[20] J. A. Clark, *Uniformly convex spaces*, Trans. Amer. Math. Soc. **40** (3) (1936), 396—414

[21] G. Da Prato, J. Zabczyk, *Convergence to equilibrium for classical and quantum spin systems*, Probab. Theory Related Fields **103** (1995), no. 4, 529-–552

[22] G. Da Prato, J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press, London Mathematical Society Lecture note series 229, 1996

[23] G. Da Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, 1992

[24] J. L. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. **63** (1948), 394–421

[25] F. Dragoni, V. Kontis, B. Zegarliński, *Ergodicity of Markov Semigroups with Hörmander Type Generators in Infinite Dimensions*, Potential Analysis, October 2012, Volume **37**, Issue 3, 199–227

[26] E. B. Dynkin, *Markov Processes*, Vol. I, II, Springer-Verlag, 1965

[27] E. B. Dynkin, *Theory of Markov processes*, Prentice-Hall, Englewood Cliffs, New Jersey, 1961 (Original Russian ed., Moscow, 1959)

[28] J. Fritz, *Infinite lattice systems of interacting diffusion processes, existence and regularity properties.* Z. Wahrsch. Verw. Gebiete **59** (1982), no. 3, 291–309

[29] L. Goldstein, *Minimizing noisy functionals in Hilbert space: An extension of the Kiefer-Wolfowitz procedure*, J. Theoret. Probab. **1** (1988), 189–204

[30] A. Guionnet, B. Zegarliński, *Lectures on logarithmic Sobolev inequalities*, Séminaire de Probabilités, **XXXVI**, 1-–134, Lecture Notes in Math., 1801, Springer, Berlin, 2003

[31] I. Gyöngy, T. Pröhle, *On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem*, Comput. Math. Appl. **19** (1990), no. 1, 65-–70

[32] M. Hairer, *Convergence of Markov processes*, Lecture notes, see http://www.hairer.org/notes/Convergence.pdf

[33] M. Hairer, J. C. Mattingly, *Yet another look at Harris' ergodic theorem for Markov chains*, Seminar on Stochastic Analysis, Random Fields and Applications VI Progress in Probability Volume **63** (2011), 109–117

[34] M. Hofmanová, J. Seidler, *On weak solutions of stochastic differential equations*, Stoch. Anal. Appl. **30** (2012), no. 1, 100–121

[35] R. Holley, D. Stroock *Diffusions on an infinite dimensional torus*, J. Funct. Anal. **42** (1981), no. 1, 29–63

[36] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171

[37] K. Ichihara, H. Kunita, *A classification of the second order degenerate elliptic operators and its probabilistic characterization*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **30** (1974), 235–254

[38] R. Khasminskii, *Stochastic stability of differential equations*, Springer-Verlag, Berlin Heidelberg 2012 (Originally published in Russian by Nauka, Moscow 1969)

[39] J. Kiefer, J. Wolfowitz, *Stochastic estimation of the maximum of a regression function*, Ann. Math. Statist. **23** (1952), 462–466

[40] C. Kipnis, C. Landim, *Scalling limits of interacting particle systems*, Springer - Verlag, Berlin, 1999

[41] V. Kontis, M. Ottobre, B. Zegarliński, *Markov semigroups with hypocoercive - type generator in infinite dimensions: Ergodicity and smoothing*, J. Funct. Anal. **270** (2016), 3173–3223

[42] V. Kontis, M. Ottobre, B. Zegarliński, *Markov semigroups with hypocoercive - type generator in infinite dimensions II: Applications*, preprint 2013, arXiv:1306.6453 [math-ph]

[43] A. P. Korostelev, *Stochastic recurrent procedures. Local properties* (Russian), Nauka, Moscow 1984

[44] N. V. Krylov, *A relatively short proof of Itô's formula for SPDEs and its applications*, Stoch. Partial Differ. Equ. Anal. Comput. **1** (2013), 152–174

[45] N. V. Krylov, B. L. Rozovskiĭ, *Stochastic evolution equations* (Russian), Current problems in mathematics, Vol. **14**, pp. 71–147, VINITI, Moscow 1979 (English translation: *J. Soviet Math. 16* (1981), 1233-1277)

[46] H. J. Kushner, D. S. Clark, *Stochastic approximation methods for constrained and unconstrained systems*, Springer, New York 1978

[47] H. J. Kushner, A. Shwartz, *Stochastic approximation in Hilbert space: Identification and optimization of linear continuous parameter systems*, SIAM J. Control Optim. **23** (1985), 774–793

[48] H. J. Kushner, G. G. Yin, *Stochastic approximation and recursive algorithms and applications*, 2nd ed., Springer, New York 2003

[49] N. Lazrieva, T. Sharia, T. Toronjadze, *The Robbins-Monro type stochastic differential equations. I. Convergence of solutions*, Stochastics Stochastics Rep. **61** (1997), 67–87

[50] N. Lazrieva, T. Sharia, T. Toronjadze, *The Robbins-Monro type stochastic differential equations. II. Asymptotic behaviour of solutions*, Stoch. Stoch. Rep. **75** (2003), 153–180

[51] N. Lazrieva, T. Sharia, T. Toronjadze, *Semimartingale stochastic approximation procedure and recursive estimation*, J. Math. Sci. (N.Y.) **153** (2008), 211–261

[52] N. Lazrieva, T. Toronjadze, *The Robbins-Monro type stochastic differential equations. III. Polyak's averaging*, Stochastics **82** (2010), 165–188

[53] T. Liggett, *Interacting particle systems*, Springer - Verlag, New York, 1985

[54] W. Liu, M. Röckner, *Stochastic partial differential equations: An introduction*, Springer, Cham 2015

[55] J. Mattingly, A. Stuart, D. Higham, *Ergodicity for SDEs and Aproximations : Locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl. **101** (2002), no. 2, 185–232

[56] S. Meyn, R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, 1993

[57] S. Meyn, R. L. Tweedie, *Stability of Markovian processes. II. Continuous-time processes and sampled chains*, Adv. in Appl. Probab. **25** (1993), no. 3, 487–517

[58] S. Meyn, R. L. Tweedie, *Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes*, Adv. in Appl. Probab. **25** (1993), no. 3, 518–548

[59] J. Munkres, *Topology*, Prentice Hall, Inc., 2nd edition, 2000

[60] M. Mureşan, A concrete approach to classical analysis, Springer, New York, 2009. xviii+433 pp.

[61] M. B. Nevel'son, R. Z. Khas'minskiĭ, *Stochastic approximation and recursive estimation*, American Mathematical Society, Providence 1973)

[62] R. Nixdorf, *An invariance principle for a finite dimensional stochastic approximation method in a Hilbert space*, J. Multivariate Anal. **15** (1984), 252–260

[63] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer - Verlag Berlin Heidelberg, 2nd edition, 2006

[64] B. Øksendal, *Stochastic differential equations. An introduction with applications*, Sixth edition, Springer - Verlag, Berlin, 2003

[65] E. Pardoux, *Équations aux dérivées partielles stochastiques non linéaires monotones. Etude de solutions fortes de type Ito*, Thèse, Université Paris Sud, 1975, see https://www.i2m.univ-amu.fr/~pardoux/Pardoux_these.pdf

[66] G. Pflug, *Stetige stochastische Approximation*, Metrika **26** (1979), 139–150

[67] R. Precup, *Linear and semilinear partial differential equations*, Walter de Gruyter & Co., Berlin 2013

[68] C. Prevot, M. Röckner, *A Concise course on Stochastic Partial differential equations*, Springer LN in Math. 1905, Berlin 2007

[69] J. Ren, M. Röckner, F. Wang, *Stochastic generalized porous media and fast diffusion equations*, J. Differ. Equations **238** (2007), 118–152

[70] D. Revuz, M. Yor, *Continuous martingales and Brownian motion*, 3rd ed., Springer-Verlag, 1999

[71] L. Rey-Bellet, *Ergodic properties of Markov processes*, Open quantum systems. **II**, 1-–39, Lecture Notes in Math., 1881, Springer, Berlin, 2006

[72] L. Rogers, D. Williams, *Diffusions, Markov Processes, and Martingales*,vol I, II, John Wiley & Sons, 1987

[73] J. Seidler, F. Zak, *A note on Continuous-time stochastic approximations in Infinite Dimensions*, preprint 2016

[74] J. Seidler, *Ergodic behaviour of stochastic parabolic equations*, Czechoslovak Mathematical Journal **47** (1997), issue 2, 277–316

[75] J. Simon, *Régularité de la solution d'une équation non linéaire dans* $\mathbb{R}^N$, Journées d'Analyse Non Linéaire (Besançon, 1977), pp. 205–227, Lecture Notes in Math. 665, Springer, Berlin 1978

[76] P. Sousi, *Advanced probability*, Lecture notes, see http://www.statslab.cam.ac.uk/~ps422/teaching.html

[77] D. Stroock, S. Varadhan, *Multidimensional Diffusion Processes*, Springer - Verlag, Berlin and New York, 1979

[78] D. Stroock, S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth

Berkeley Symposium on Mathematical Statistics and Probability 1972, Vol. **III**: Probability theory, 333–359

[79] G. Yin, *On H-valued stochastic approximation: Finite dimensional projections*, Stochastic Anal. and Appl. **10** (1992), 363–377

[80] G. Yin, Y. M. Zhu, *On H-valued Robbins-Monro processes*, J. Multivariate Anal. **34** (1990), 116–140

[81] L. Zambotti, *An analytic approach to existence and uniqueness for martingale problems in infinite dimensions*, Probab. Theory Related Fields **118** (2000), no. 2, 147–168

[82] B. Zegarliński, *The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice*, Communications in Mathematical Physics, January (III) 1996, Volume **175**, Issue 2, 401–432

[83] F. Zak, *Existence of equilibrium for infinite system of interacting diffusions*, Stochastic Anal. and Appl. **34** (2016), 1057–1082