Common noise pullback attractors for stochastic dynamical systems.

Federico Graceffa and Jeroen S.W. Lamb

Department of Mathematics,
Imperial College London,
180 Queen’s Gate,
London SW7 2AZ,
United Kingdom

12th August 2021

Abstract

We consider SDEs driven by two different sources of additive noise, which we refer to as intrinsic and common. We establish almost sure existence and uniqueness of pullback attractors with respect to realisations of the common noise only. These common noise pullback attractors are smooth probability densities that depend only on (the past of) a common noise realisation and to which the pullback evolution of a corresponding stochastic Fokker-Planck equation converges. Common noise pullback attractors have a natural motivation in the context of particle systems with intrinsic and common noise, describing the distribution of the system conditioned on (the past of) a common noise realisation.
1 Introduction and summary of the main results

1.1 Motivation

In the theory of dynamical systems, broadly speaking, two dominant points of view may be distinguished: the topological point of view (understanding of the dynamics at the level of (typical) individual trajectories) and the probabilistic point of view (understanding of the dynamics at the level of average statistical properties, e.g. through Ergodic Theory) [34]. Dynamical systems in the presence of noise (such as stochastic dynamical systems defined by SDEs) are predominantly approached from the latter point of view, with powerful analytical techniques from stochastic analysis and Markov processes [40]. The alternative *random dynamical systems* approach considers dynamical systems with noise as skew-product systems,
where noise drives an otherwise deterministic dynamical system and the noise driving process admits a pathwise and probabilistic (in terms of ergodic theory) description. The latter allows a blend of the traditional topological and probabilistic approaches to achieve probabilistic results about the behaviour of trajectories of the (non-autonomous) noise-driven system. For instance, Arnold and co-workers [2] have established the existence of random generalisations of attractors, as well as stable, unstable and centre manifolds.

In this paper we develop a random dynamical systems point of view for SDEs with two distinguished sources of noise, which we refer to as intrinsic and common in view of the motivating example of a system with identical non-interacting particles (or agents) that are subject to intrinsic noise at the level of each particle and a common noise that is equal to all. Such settings naturally arise in a broad range of applications, for instance in genetics [19, 23, 35, 36], neuroscience [14, 38], epidemics [32], pattern formation [44] and financial mathematics [31].

The aim of this paper is to study a stochastic dynamical system with intrinsic and common noise, conditioned on the past of the common noise realisation. In the context of the motivating particle system, this yields a description of the probability distribution of the particle system, the evolution of which is described by a stochastic Fokker-Planck equation, subject to the past of the realisation of the common noise. We establish the existence and uniqueness of a corresponding common noise pullback attractor for a specific class of SDEs where the intrinsic and common noises are additive Brownian motions. Our approach is not limited to this special class, but the main aim of this paper is to develop the concept of common noise pullback attractors in this specific transparent setting.

Common noise pullback attractors facilitate the study of time-averaged properties of the distribution describing the particle system, through the application of ergodic theory, cf. equation (1.5) in Section 1.2 for example observing the variance of the distribution as a measure of synchronisation. Synchronisation is a widely studied dynamical phenomenon in complex systems with ramifications in a wide range of applications [43]. In addition, from a modelling perspective, our point of view is natural where the intrinsic noise is inherently or practically not observable, while the common noise can in principle be observed. Examples include the sentiment of traders in markets and voters under the influence of mass media, where the latter can be treated as a stochastic process or as a deterministic signal, leading to the consideration of dynamics with common noise or more general non-autonomous dynamics. In fact, the random dynamical systems approach taken in this paper in principle allows us to address both settings at the same time.

The notion of pullback attractors is well-established in non-autonomous and random dynamical systems, see for instance [2, 36] and [3, 13, 17, 29, 39, 50, 51] in the context of SPDEs. The analysis of pullback attractors in applications of complex nonlinear systems is gaining popularity in recent years, for instance in the context of climate science and turbulence [20, 30]. This paper is a further contribution in this direction.

---

1Bressloff calls this the population or SPDE perspective, in contrast with the particle perspective, where one averages over both the intrinsic and common noise [12].
1.2 Common noise pullback attractor

We consider SDEs in $\mathbb{R}^d$ of the form

$$dx(t) = -\nabla V(x(t))dt + \sigma dW(t) + \eta dB(t), \quad (1.1)$$

where $V$ represents a smooth potential, $\sigma, \eta$ are positive definite matrices, $W$ is a Brownian motion and $B$ represents another source of noise, such as another Brownian motion or related process.

We refer to $W$ as intrinsic and $B$ as common noise, motivated by the setting of a system of identical particles with states $x_i \in \mathbb{R}^d$, $i \in \{1, \ldots, N\}$, each subject to an intrinsic noise $W_i$ and an identical common noise $B$

$$dx_i(t) = -\nabla V(x_i(t))dt + \sigma dW_i(t) + \eta dB(t). \quad (1.2)$$

In the limit of large $N$, the evolution of this particle system is described by the evolution of a measure on $\mathbb{R}^d$. This measure has a Lebesgue density $p$ whose evolution is governed by the stochastic Fokker-Planck equation

$$dp = \left[ \Delta V(x)p + \nabla V(x)\frac{\partial p}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 p}{\partial x^2} \right] dt - \eta \frac{\partial p}{\partial x} \circ dB(t), \quad (1.3)$$

where, as usual, $\circ$ refers to the Stratonovich convention of stochastic integration.

Similar equations have been derived earlier by Giles and Reisinger [31] in the context of pricing baskets of financial derivatives, Bressloff [12] in the context of neuronal dynamics, Bain and Crisan [4] in the context of stochastic filtering and Carmona and Delarue [14] in the context of mean field games.

We aim to employ (1.3), like Bressloff [12], to study the evolution of population densities as a function of the common noise. Thereto, we approach the stochastic Fokker-Planck equation (1.3) from a random dynamical systems point of view, considering the non-autonomous evolution of the density $p$ as a function of the realisation $\beta$ of the common noise $B$. In Section 2 and in particular Theorem 2.2, we provide a detailed discussion on the existence, uniqueness and regularity of solutions to the initial value problem of (1.3) as a deterministic non-autonomous PDE, for a sufficiently regular common noise realisation. It is shown that the flow evolves initial conditions in $L^1$ to the Schwartz space of smooth rapidly decaying Lebesgue densities. In Section 4, we establish that the SPDE (1.3) is a random dynamical system in the sense that it admits a description in skew-product form with ergodic base dynamics generating the common noise $B(t)$, see Lemma 4.1.

Traditionally, dynamical systems theory focuses mostly on the long-term behaviour of solutions. In the non-autonomous setting, as it is rare to have convergence in forward time (since the equations of motion vary with time), it is natural to consider the asymptotic behaviour of pullback dynamics instead, which has better prospects of convergence and reveals important aspects of the dynamics. Let $\Phi(t, \beta)$ represent the time-$t$ flow of (1.3) with common noise realisation $\beta$. Instead of studying the behaviour of initial conditions

\footnote{Often, (1.3) is written in Itô form, resulting in an additional term $\eta^2$ added to $\sigma^2$ in the diffusion term.}
with fixed noise realisation in the limit $t \to \infty$, pullback dynamics considers the behaviour of initial conditions of this flow, fixing the end-time, say at $t = 0$, in the limit of the starting time $\tau \to -\infty$. A pullback attractor describes the state of the system, conditioned on the past of the time-dependent input (noise realisation). In the context of particle dynamics with intrinsic and common noise, the objective is to describe the distribution of the particle system with intrinsic noise, subject to the past realisation of the common noise.

Under a natural assumption on the potential $V$ that guarantees the existence of a unique stationary density for (1.1) in the absence of common noise, the main result of this paper is that the stochastic Fokker-Planck equation (1.3) has a unique pullback attractor that is a random equilibrium, i.e. for almost all common noise realisations $\beta$, the limit

$$p_\beta := \lim_{\tau \to -\infty} \mathcal{W} (\tau, \theta_{-\tau} \beta) p$$

(1.4)

exists in the Schwartz space and is independent of the initial probability density $p \in L^1$, see Theorem 4.2. This result relies on the fact, obtained in Section 3, that for almost every noise realisation the non-autonomous evolution is a contraction. Moreover, the convergence is uniform in the common noise realisation. We refer to $p_\beta$ as the common noise pullback attractor of the SDE (1.1). It turns out that $p_\beta$ is the density of the measure obtained by averaging, for a fixed common noise realisation $\beta$, the canonical $(\omega, \beta)$-dependent pullback measures of SDE (1.1) over all intrinsic noise relations $\omega$, cf. Proposition 4.4.

From a random dynamical systems point of view, $p_\beta$ in (1.4) is called a globally attracting random equilibrium of the stochastic Fokker-Planck equation (1.3). This is the simplest type of attractor one may encounter in a random dynamical system. In general, random (pullback) attractors may display more complicated behaviour, cf. [17].

Finally, by virtue of ergodicity we find (in Proposition 4.5) that if $g$ is a continuous observable on the relevant solution space of densities for (1.3), $\mathbb{P}_B$-almost surely,

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau g(\Phi(\tau, \theta_{-\tau} \beta) p) dt = \mathbb{E}^{\mathbb{P}_B}[g(p_{\cdot})],$$

(1.5)

with $\mathbb{E}^{\mathbb{P}_B}$ denoting the expectation with respect to the probability measure $\mathbb{P}_B$ of the common noise. For special types of observables, the expectation (1.5) is related to an expectation with respect to the stationary measure $\rho$ of the SDE (1.1). In particular, when the observable $g$ is a $p_\beta$-expectation of a continuous observable $h : \mathbb{R}^d \to \mathbb{R}$, i.e. $g(p_\beta) = \int_{\mathbb{R}^d} h(x)p_\beta(x) dx$, then

$$\mathbb{E}^{\mathbb{P}_B}[g(p_{\cdot})] := \int_{\Omega_B} g(p_\beta) \mathbb{P}_B(d\beta) = \int_{\Omega_B} \int_{\mathbb{R}^d} h(x)p_\beta(x) dx \mathbb{P}_B(d\beta)$$

(1.6)

$$= \int_{\mathbb{R}^d} h(x) \int_{\Omega_B} p_\beta(x) \mathbb{P}_B(d\beta) dx = \int_{\mathbb{R}^d} h(x) \rho(dx) =: \mathbb{E}^{\rho}[h]$$

However, in general the expectation (1.5) is not expressible in terms of the stationary measure $\rho$ of (1.1). For instance, the variance

$$\text{Var}(p_\beta) := \mathbb{E}^{\rho_\beta}[x^2] - (\mathbb{E}^{\rho_\beta}[x])^2$$

which is an indicator of synchronisation (of the particle system), is an observable whose $\mathbb{P}_B$-expectation cannot be deduced from $\rho$, cf. the examples discussed in Section 1.3.
1.3 Examples

1.3.1 Ornstein-Uhlenbeck process with intrinsic and common noise

Our results are well-illustrated in the elementary example of an Ornstein-Uhlenbeck process with intrinsic and common noise

\[ dx(t) = -ax(t)dt + \sigma dW(t) + \eta dB(t), \]  

(1.7)

with \( x \in \mathbb{R}, a > 0 \) and \( W \) and \( B \) are independent Brownian motions.\(^3\) Due to the linearity of (1.7) the solution of (1.3) with initial condition \( \delta_x(s) \) and common noise realisation \( \beta \) can be explicitly calculated (for all \( t > s \)) to have the form

\[
p(x, t) = \frac{a}{\sqrt{\pi \sigma^2 (1 - e^{-2a(t-s)})}} \exp \left( -\frac{a}{\sigma^2 (1 - e^{-2a(t-s)})} (x - m_\beta(t, s))^2 \right),
\]

where

\[
m_\beta(t, s) := x(s)e^{-a(t-s)} + \eta \int_s^t e^{-a(t-u)}d\beta(u),
\]

and the latter integral is \( \mathbb{P}_B \)-almost surely finite. Indeed, by Theorem 4.2, the unique common noise pullback attractor of (1.7) is independent of the initial condition (in \( L^1 \), cf. Section 2) and thus equals \( \mathbb{P}_B \)-almost surely the normal distribution

\[
p_\beta(x) = \lim_{s \to -\infty} p(x, 0) = \frac{a}{\sqrt{\pi \sigma^2}} \exp \left( -\frac{a}{\sigma^2} \left( x - \eta \int_{-\infty}^0 e^{au}d\beta(u) \right)^2 \right).
\]

(1.8)

This example illustrates how the exact synchronisation of solutions of (1.7) in the absence of intrinsic noise (\( \sigma = 0 \)) turns into an approximate synchronisation of the corresponding particle system in the presence of small intrinsic noise (\( \sigma \ll 1 \), characterized by small \( \text{Var}(p_\beta) \)). Namely, in the absence of intrinsic noise, \( \mathbb{P}_B \)-almost surely all pairs of initial conditions \( x, y \in \mathbb{R}^d \) pathwise converge, i.e. \( x_\beta(t), y_\beta(t) \) of (1.7) with noise realisation \( \beta \) satisfy \( \lim_{t \to \infty} |x_\beta(t) - y_\beta(t)| = 0 \) \(^4\), while in the presence of intrinsic noise the distribution converges to a normal distribution with variance \( \text{Var}(p_\beta) = \frac{\sigma^2}{2a} \). The location of this normal distribution depends on (the past of) the common noise realisation \( \beta \), i.e. the mean \( m_\beta) = m_\beta(0, -\infty) \) and is independent of the intrinsic noise strength \( \sigma \). In view of (1.5), this implies

\[
\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \text{Var}(\Phi(\tau, \beta)p)dt = \frac{\sigma^2}{2a} \quad \text{and} \quad \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau m(\Phi(\tau, \beta)p)dt = \mathbb{E}(x) = 0,
\]

where \( \rho \) denotes the stationary measure of (1.7).

We contrast the average of the observed variance along trajectories of (1.3) with the fact that the stationary density \( p_\rho \) of (1.7), \( p_\rho(x) = \int p_\beta(x)\mathbb{P}(d\beta) \), has a different variance. In particular,

\[
p_\rho(x) = \sqrt{\frac{a}{\pi \sigma^2 \eta^2}} \exp \left( -\frac{a}{(\eta^2 + \sigma^2)} x^2 \right)
\]

\(^3\)Most of the results here do not require \( B \) to be a Brownian motion, cf. also footnote \(^9\).

\(^4\)For details, see Appendix D.
is a normal distribution with mean 0 and variance $\frac{\sigma^2 + \eta^2}{2a}$. Indeed, synchronisation of the particle system corresponds to localisation of the pullback measure, rather than to localisation of the stationary measure. If $\sigma$ is small and $\eta$ is large, the particle distribution is asymptotically strongly localized, while the stationary distribution is not.

We note that the Ornstein-Uhlenbeck example (1.7) is very special, in particular the fact that the shape of the density $p_\beta$ does not depend on the noise realisation $\beta$. This is a consequence of the linearity of this example, which also yields it exactly solvable.

### 1.3.2 Bi-stable dynamics with intrinsic and common noise

We next consider the less degenerate, nonlinear, example of (1.1) with $x \in \mathbb{R}$ and $V(x) = \frac{1}{4}x^4 - \frac{a^2}{2}x^2$ is a double-well potential

$$dx(t) = x(t)(a - x(t)^2)dt + \sigma dW(t) + \eta dB(t), \quad (1.9)$$

with $a > 0$ and $\sigma, \eta$ constants as above and . $W(t)$, $B(t)$ are Brownian motions. In this case, the stationary probability density of the SDE (1.9) admits the explicit expression

$$p_\rho(x) = \frac{1}{N} \exp \left( -\frac{2}{\sigma^2 + \eta^2} \left( \frac{x^4}{4} - ax^2 \right) \right), \quad (1.10)$$

where $N := \int \exp \left( -\frac{2}{\sigma^2 + \eta^2} \left( \frac{x^4}{4} - ax^2 \right) \right) dx$ is a normalization constant

An important difference with (1.7) is that (1.9) is nonlinear. To our best knowledge, in this case, the common noise pullback attractor $p_\beta$ of (1.9) does not admit a comprehensive analytical expression, but it can be approximated numerically (for instance, by means of Monte Carlo methods, cf. [35]). In Figure 1 some numerically obtained examples of densities for common noise pullback attractors are presented, illustrating how the stationary density (depicted in the background in grey) may differ substantially from the densities of individual pullback attractors $p_\beta$ which depend on the common noise realisation $\beta$. This figure illustrates some of the limitations in dynamical information that a stationary measure of a stochastic dynamical system provides.

There are various important questions concerning common noise pullback attractors that we have not addressed here, but which deserve further attention. For instance, it would be of interest to determine the support of the stationary measure of (1.3), i.e. the range of possible densities of common noise pullback attractors, in particular also as a function of system parameters. In the exactly solvable Ornstein-Uhlenbeck example of Section 1.3.1 the range consists of a one-parameter family of normal distributions, with identical variance depending on the strength of the intrinsic noise. In the double-well example of Section 1.3.2 Figure 1 suggests that the range is also limited but with a more complicated dependence on system parameters.
Figure 1: Densities of common noise pullback attractors $p_\beta$ of the SDE with double well potential with intrinsic and common additive noise (1.9), with $a = 1, \sigma^2 + \eta^2 = 1$ and (a) $\sigma \ll \eta = 0.99$, (b) $\sigma = \eta = \frac{1}{2}\sqrt{2}$ and (c) $\sigma \gg \eta = 0.15$. Pullback attractors for different common noise realisations $\beta$ are represented by graphs with different colors. The stationary density $p_\rho$ (1.10) of the SDE (1.9), plotted in grey in the background, is identical in all three cases. Different scales on the $p_\beta$-axis have been chosen so as to achieve similar resolutions in the graphs of the densities of the common noise pullback attractors. When common noise dominates intrinsic noise, $\sigma \ll \eta$ (a), one predominantly observes localized pullback attractors, corresponding to approximate synchronisation. When intrinsic noise dominates common noise, $\sigma \gg \eta$ (c), the densities of the common noise pullback attractors tend to be less localized and relatively close to the stationary density. To obtain an objective quantification of the degree of synchronisation, we have numerically approximated the time-averaged variance of (particle) distributions, as $E^\beta [\text{Var}(p)]$ by virtue of (1.5), yielding the values (a) 0.04, (b) 0.53 and (c) 0.90 (the latter being close to $\text{Var}(p_\rho)$), in accordance with the perceived degrees of localisation in the density graphs.
2 The non-autonomous Fokker Planck equation and its initial value problem

In this section we consider the derivation and analysis of equation (1.3), as a non-autonomous Fokker-Planck equation. This forms the basis of our discussion of (1.3) in Section 4 in the stochastic setting, as a random dynamical system.

In Section 2.1, we discuss the derivation of the non-autonomous Fokker-Planck equation (1.3) from two points of view: as the Fokker-Planck equation for a non-autonomous SDE and from a particle system approximation, which motivates the choice of terminology common noise. In Section 2.2 we establish the existence and uniqueness of the solution for the non-autonomous Fokker-Planck equation (1.3) within a suitable setting and discuss how this solution smoothens when \( t > 0 \), given an initial condition \( q_0 \in L^1(\mathbb{R}^d) \) at \( t = 0 \). The techniques employed are in principle deterministic and well-established, see e.g. [16, 25, 47], but as our specific non-autonomous setting is not normally addressed, we present a self-contained technical discussion in Appendix A. The choice of initial conditions in \( L^1(\mathbb{R}^d) \) rather than in \( L^2(\mathbb{R}^d) \), as is commonly found in the literature, is to cater for natural densities relevant to the particle system interpretation, such as Dirac’s delta, representing a system with all particles in the same initial state. It turns out that initial conditions in \( L^1(\mathbb{R}^d) \) evolve immediately into \( L^2(\mathbb{R}^d) \).

2.1 Derivation of the non-autonomous Fokker-Planck equation

2.1.1 The Fokker-Planck equation of the non-autonomous SDE

Let us consider (1.1) as a non-autonomous SDE where \( B(t) = \beta(t) \) is deterministic. We let \( \beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d) \), defined as the set of functions from \( \mathbb{R} \) to \( \mathbb{R}^d \) that are everywhere locally \( \alpha \)-Hölder continuous for any \( \alpha < \frac{1}{2} \). Writing \( y(t) := x(t) - \eta \beta(t) \), the SDE (1.1) can be written as

\[
dy(t) = -\nabla V(y(t) + \eta \beta(t))dt + \sigma dW(t).
\]

With \( U(y, t; \beta) := V(y(t) + \eta \beta(t)) \), we find

\[
\nabla U(y, t; \beta) = \nabla V(y + \eta \beta(t)), \quad \Delta U(y, t; \beta) = \Delta V(y + \eta \beta(t))
\]

where \( \nabla \) and \( \Delta \) denote the gradient and Laplacian with respect to the first argument. The Fokker-Planck equation describing the annealed evolution of Lebesgue probability densities \( q \) associated with the SDE (2.1) is given by

\[
\partial_t q = \nabla (\nabla U(y, t; \beta) q) + \frac{1}{2} \sigma^2 \Delta q.
\]

Transforming variables \( y \) back to \( x \) in (2.2), with densities \( p(x) := q(y) \), yields (1.3).

\[\text{[5]}\text{The choice of this regularity is motivated by the fact that pathwise realisations of the Brownian motion possess this regularity almost surely, see e.g. [33].}\]
2.1.2 The Fokker-Planck equation of the non-autonomous particle system

The Fokker-Planck equation (1.3) can also be motivated directly from a particle system point of view, following for instance [37]. Let us consider a system of particles $x_i$, $i = 1, \ldots, N$ satisfying (1.2)

$$dx_i(t) = -\nabla V(x_i(t))dt + \sigma dW_i(t) + \eta d\beta(t),$$

with $W_i$ independent Brownian motions representing the intrinsic noise and $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$ a deterministic common driving. Let $\varphi \in C_b^2(\mathbb{R}^d)$ be an observable which is a bounded function of compact support with bounded first and second derivative. Its evolution is given by

$$\varphi(x_i(t)) = \varphi(x_i(0)) + \int_0^t \left(-\nabla^T \varphi(x_i(s)) \nabla V(x_i(s)) + \frac{1}{2} \sigma^2 \Delta \varphi(x_i(s))\right) ds$$

$$+ \int_0^t \sigma \nabla^T \varphi(x_i(s)) dW_i(s) + \int_0^t \eta \nabla^T \varphi(x_i(s)) d\beta(s).$$

(2.3)

It is crucial to recognise that the stochastic integral with respect to the intrinsic noise $W_i$ represents a distribution, while the other integrals yield scalars. The empirical measure for a particle distribution may be defined, as usual, as

$$\nu(t) := \lim_{N \to +\infty} \frac{1}{N} \sum_{i=1}^N \delta_{x_i(t)},$$

where $\delta_x$ denotes the Dirac measure at $x$ and we consider convergence in the weak topology of the collection of all finite signed Borel probability measures on $\mathbb{R}^d$. Integrating both sides of (2.3) with respect to this empirical measure, denoting $\nu_t(\varphi) := \int \varphi d\nu_t$, yields the SPDE

$$\nu_t(\varphi) = \nu_0(\varphi) + \int_0^t \nu_s(A_1 \varphi) ds + \int_0^t \nu_s(A_2 \varphi) d\beta(s),$$

(2.4)

where $A_1, A_2$ are the differential operators

$$A_1 \varphi := \frac{1}{2} \sigma^2 \Delta \varphi - \nabla^T \varphi \nabla V$$

$$A_2 \varphi := \eta \nabla^T \varphi.$$

In particular, the term containing the stochastic integral vanishes, see [37, Proof of Theorem 3.1]. Following [4, Chapter 7.3], assuming that the empirical measure $\nu$ has a sufficiently smooth Lebesgue density $p(t)$, one may reformulate (2.4) as

$$\nu_t(\varphi) = \int_\mathbb{R} \varphi(x) p(t, x) dx$$

$$= \int_\mathbb{R} \varphi(x) \left(p(0, x) + \int_0^t A_1^* p(s, x) ds + \int_0^t A_2^* p(s, x) d\beta(s)\right) dx,$$

(2.5)

The Glivenko-Cantelli theorem [48] provides more detail on the convergence in the weak topology of the empirical measure. See Boissard and Le Gouic [10], for a discussion of the convergence of the empirical measure with respect to the Wasserstein distance.
where $A_1^*$, $A_2^*$ are given by

\[
\begin{align*}
A_1^* \psi &= \frac{1}{2} \sigma^2 \Delta \psi + \nabla (\nabla V \psi) = \frac{1}{2} \sigma^2 \Delta \psi + \Delta V \psi + \nabla \nabla V \\
A_2^* \psi &= -\eta \nabla^T \psi.
\end{align*}
\]

We thus obtain

\[
p(t, x) = p(0, x) + \int_0^t A_1^* p(s, x) ds + \int_0^t A_2^* p(s, x) d\beta(s),
\]

which is the integral form of [1,3].

### 2.2 The initial value problem: existence, uniqueness and regularity

We denote by $C^\infty_c(\mathbb{R}^d)$ the space of all smooth and compactly supported functions on $\mathbb{R}^d$ and

\[
\langle q, \varphi \rangle := \int_{\mathbb{R}^d} \varphi q dx.
\]

We further denote by $\mathcal{P}(\mathbb{R}^d)$ the space of all Borel probability measures on $\mathbb{R}^d$. Our notion of a weak solution is as follows:

**Definition 2.1 (Weak $L^1$-probability solution).** A function $q \in C([0, T]; L^1(\mathbb{R}^d); C^{1/2}(\mathbb{R}, \mathbb{R}^d))$ is a weak $L^1$-probability solution of the initial value problem for the non-autonomous Fokker-Planck equation (2.2) if $q$ solves this equation with initial condition $q_0 \in \mathcal{P}(\mathbb{R}^d)$ and $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$, such that $q(t) \in \mathcal{P}(\mathbb{R}^d)$ for all $t > 0$ and $\langle q(t), \varphi \rangle \to \langle q_0, \varphi \rangle$ as $t \to 0$ for all $\varphi \in C^\infty_c(\mathbb{R}^d)$.

Our main result establishes the existence and uniqueness of probability solutions of the $L^1$ initial value problem and the fact that such solutions are smooth and their derivatives of any order are rapidly decreasing after any finite time, i.e. they belong to the Schwartz space

\[
S := S(\mathbb{R}^d) := \left\{ f \in C^\infty(\mathbb{R}^d) : \forall n, m \in \mathbb{N}^d, \| f \|_{n, m} < \infty \right\},
\]

where $C^\infty(\mathbb{R}^d)$ denotes the space of infinitely differentiable functions on $\mathbb{R}^d$ and

\[
\| f \|_{n, m} := \sup_{x \in \mathbb{R}^d} |x^n (D^m f)(x)|.
\]

**Theorem 2.2 (Existence, uniqueness and regularity).** Let $q_0 \in \mathcal{P}(\mathbb{R}^d)$ and $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$. Let us assume the potential $U$ is $C^\infty$ in $x$ and satisfies the dissipation condition

\[
\nabla U(x, t; \beta) \cdot x \| x \|^2 \geq \frac{1}{2} \| x \|^6 - C|\beta|^6
\]

for some constant $C > 0$. Then, the non-autonomous Fokker-Planck equation (2.2) with initial condition $q(0) = q_0$ admits a unique weak probability solution $q(t) \in \mathcal{P}(\mathbb{R}^d)$ for all $t > 0$ which is everywhere locally $\alpha$-Hölder continuous in time for any $\alpha < \frac{1}{2}$, such that $q(\cdot, t; \beta) \in S$ for $t > 0$.

---

7Here we use the multi-index notation, as in Evans [26].
We defer the proof of this result to Appendix A.

It turns out that, $L^1$ solutions are not unique in general. In Appendix A, it is shown that uniqueness is ensured by a weighted integrability condition, expressing the fact that no mass can come from infinity, nor disappears through infinity, in finite time. It turns out (Lemma A.1) that this weighted integrability condition is equivalent to the conservation of mass, hence ensuring existence and uniqueness in the context of probability solutions. In view of the particle systems motivation of Section 2.1.2, this setting is natural as it concerns the conservation of particles.

Remark 2.3. Note that the dissipation condition (2.6) also ensures the existence of a unique stationary solution of (1.1). This condition is fulfilled, for example, in the case of the double well potential $-\nabla V(x) = x(a - \|x\|^2)$, cf. the example discussed in Section 1.3.2.

3 Contraction property of the non-autonomous Fokker-Planck equation

In this section we establish that the time-$t$ evolution operator $\Phi(t, \beta)$ of the non-autonomous Fokker-Planck equation (2.2) is a contraction for all $t > 0$ and $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$. Technical proofs are deferred to Appendix B.

In the autonomous setting ($\eta = 0$), if $V$ is strictly convex, $\Phi$ is known to be a contraction for all $t > 0$ if the metric on the solution space is chosen to be the usual Wasserstein distance $W^p$ for any $p \geq 1$. Moreover, strict convexity of $V$ is also a necessary condition. Under the milder assumption that the potential $V$ is strictly convex outside a given ball in $\mathbb{R}^d$, with is a less restrictive and more realistic condition, Eberle [22] established contractivity of the evolution (again in the autonomous setting) for an appropriately chosen Kantorovich-Rubinstein metric. In this section, we adapt the results from [22] for autonomous Fokker-Planck equations to the non-autonomous setting.

Let us consider again (1.1) as a non-autonomous SDE where $B(t) = \beta(t)$ at all times $t$ for some $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$,

$$dx(t) = -\nabla V(x(t))dt + \sigma dW(t) + \eta d\beta(t), \quad (3.1)$$

where $W$ is a $d$-dimensional Brownian motion, $\sigma, \eta \in \mathbb{R}^{d \times d}$ are constant matrices with positive determinants and the potential $V$ satisfies the same assumptions as in Theorem 2.2. We denote by $\mu_{t, \beta}$ and $\nu_{t, \beta}$ the time-$t$ evolved probability measures with respect to a given input $\beta$ and initial conditions $\mu, \nu$ respectively. In other words, with $p_\mu$ denoting the Lebesgue density of $\mu$, we have

$$\mu_{t, \beta}(A) := \int_A \Phi(t, \beta)p_\mu(x)dx, \quad \forall A \in \mathcal{B}(\mathbb{R}^d), \quad t \geq 0, \quad (3.2)$$

and similarly for $\nu_{t, \beta}$. We employ a reflection coupling method to determine a bound for the distance between $\mu_{t, \beta}$ and $\nu_{t, \beta}$ with respect to some appropriately chosen metric. This method entails the introduction of an additional auxiliary process $y(t)$ such that $x(t) = y(t)$.
for \( t \geq T \), for some \( T \), adapting [22] Eqs. (2)-(3) to the non-autonomous setting:

\[
\begin{aligned}
\begin{cases}
  dy(t) = -\nabla V(y(t))dt + \sigma(I - 2e(t)e^\top(t))dW(t) + \eta d\beta(t), & \text{for } t < T \\
y(t) = x(t), & \text{for } t \geq T
\end{cases}
\end{aligned}
\]

where \( T := \inf \{ t \geq 0 : x(t) = y(t) \} \) is the coupling time and \( ee^\top \) is the orthogonal projection onto the unit vector

\[ e(t) := \frac{\sigma^{-1}(x(t) - y(t))}{|\sigma^{-1}(x(t) - y(t))|}. \]

The general aim is to construct a function \( f \) such that the process \( e^{ct}f(|x(t) - y(t)|) \) is a (local) supermartingale for \( t < T \), with a constant \( c > 0 \) that is maximized by choosing \( f \) appropriately. This ensures uniform, exponential contraction with respect to a Kantorovich-Rubinstein metric \( W_f \).

**Definition 3.1** (Kantorovich-Rubinstein distance). Let \( f \in C^2([0, \infty)) \) be concave and increasing with \( f(0) = 0, f'(0) = 1 \). The \( W_f \)-distance between two Borel probability measures \( \mu, \nu \in \mathcal{P}(\mathbb{R}^d) \) is defined by

\[
W_f(\mu, \nu) := \inf_{\gamma} \mathbb{E}\gamma[d_f(x, y)] = \inf_{\gamma} \int d_f(x, y)\gamma(dx, dy)
\]

with \( \mathbb{R}^d \) distance \( d_f(x, y) := f(|x - y|) \), where \( \| \cdot \| \) is a norm on \( \mathbb{R}^d \), and the infimum is taken over all couplings \( \gamma \) of \( \mu \) and \( \nu \).

In this paper, we always choose the norm to be \( \| \cdot \| = |\sigma^{-1} \cdot | \), with \( | \cdot | \) denoting the Euclidean norm in \( \mathbb{R}^d \) and \( \sigma \) the nondegenerate diffusion matrix from (3.1).

We adapt [22] Theorem 1 and Corollary 2 to obtain:

**Proposition 3.2** (Kantorovich-Rubinstein contraction). Consider the non-autonomous stochastic differential equation (3.1) and the setting of Theorem 2.2. Let \( \mu_{t, \beta}, \nu_{t, \beta} \) be time-\( t \) evolved probability measures, as defined in (3.2). Then, there exist a constant \( c > 0 \) and a convex and increasing function \( f \) such that for any \( \beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d), t > 0 \) and initial conditions \( \mu, \nu \in \mathcal{P}(\mathbb{R}^d) \),

\[ W_f(\mu_{t, \beta}, \nu_{t, \beta}) \leq e^{-ct} W_f(\mu, \nu). \]

Note that the function \( f \) in this proposition can be determined constructively. It turns out that convergence in the chosen Kantorovich-Rubinstein metric implies convergence in \( L^1 \).

**Proposition 3.3** (Convergence in \( L^1 \)). Consider the non-autonomous stochastic differential equation (3.1) and the setting of Theorem 2.2. Let \( \mu_{t, \beta} \) be the time-\( t \) evolved probability measure, as defined in (3.2). Assume that \( \beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d) \) is such that \( (\mu_{t, \beta})_{t > 0} \) is a Cauchy sequence with respect to \( W_f \). Then, the sequence \( (\mu_{t, \beta})_{t > 0} \) of the associated Lebesgue densities converges in \( L^1 \).

---

*Recall that a Borel measure \( \gamma \) on \( X \times X \) is called a coupling of Borel measures \( \mu \) and \( \nu \) on \( X \) if \( \gamma(A \times X) = \mu(A) \) and \( \gamma(X \times B) = \nu(B) \) for all \( A, B \in \mathcal{B}(X) \).*
At this point it is important to note that forward convergence, as obtained in Proposition 3.2, does not necessarily imply pullback convergence. While the contraction property ensures that all solutions approach each other as time progresses forwards, in order to guarantee pullback convergence additional conditions (on $\beta$) must be satisfied. For instance, boundedness of $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$ would suffice. In Section 4 we consider the stochastic setting of (1.3) and obtain in Theorem 4.2 pullback convergence for $\mathbb{P}_B$-almost all Brownian paths $\beta$.

4 The stochastic Fokker-Planck equation as a random dynamical system

In Sections 2 and 3, we have considered the Fokker Planck equation (1.3) as a non-autonomous PDE. In this Section we consider the stochastic setting with $B(t)$ a Brownian motion and show that the resulting stochastic Fokker-Planck equation (1.3) is a random dynamical system.

We establish almost sure pullback attractors, using the contractivity obtained in the non-autonomous analysis in Section 3, noting that sample paths of the Brownian motion $B(t)$ are $\mathbb{P}_B$-almost surely in $C^{1/2}(\mathbb{R}, \mathbb{R}^d)$.

The main results of this section concern the fact that (1.3) is a random dynamical system (Proposition 4.1) which (almost surely) possesses a unique pullback attractor (Theorem 4.2) and the correspondence between the common noise pullback attractor of (1.3) and a partial disintegration of the invariant Markov measure of (1.1) (Proposition 4.4). Technical proofs are deferred to Appendix C. We first recall briefly some preliminaries from the random dynamical system approach towards stochastic differential equations, involving the sample path space of Brownian motions [2, Chapters 1, 2 and Appendix A].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X$ be a metric space with Borel $\sigma$-algebra $\mathcal{B}(X)$. We consider the situation with two-sided continuous time $t \in \mathbb{R}$. A random dynamical system on $X$ consists of two components. The first is a metric dynamical system modelling the noise. This is a $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}, \mathcal{F})$-measurable function $\theta : \mathbb{R} \times \Omega \to \Omega$ such that

1. $\theta(0, \omega) = \omega$ and $\theta(t + s, \omega) = \theta(t, \theta(s, \omega))$ for all $t, s \in \mathbb{R}, \omega \in \Omega$,
2. the measure is preserved, i.e. $\mathbb{P}(\theta(t, A)) = \mathbb{P}(A)$ for all $t \in \mathbb{R}$ and $A \in \mathcal{F}$.

Moreover, $\theta$ is said to be ergodic if for any $A \in \mathcal{F}$, $\theta_t A = A$ for all $t \in \mathbb{R}$ implies $\mathbb{P}(A) \in \{0, 1\}$. The second component is a mapping that models the dynamics of the system. This is a $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{F} \otimes \mathcal{B}(X), \mathcal{B}(X))$-measurable function $\phi : \mathbb{R} \times \Omega \times X \to X$ such that

1. $\phi(0, \omega, x) = x$ for all $x \in X$,
2. $\phi(t + s, \omega, x) = \phi(t, \theta_s \omega, \phi(s, \omega, x))$ for all $t, s \in \mathbb{R}$ and $x \in X$, $\mathbb{P}$-almost surely (cocycle property).

Our results extend naturally to other common noise processes $B(t)$; for instance, those described by an SDE of the form $dB(t) = f(B(t))dt + \eta d\tilde{W}(t)$ for some $f \in C^1$ and Brownian motion $\tilde{W}(t)$, such as the Ornstein-Uhlenbeck process.

Here and throughout the paper we will use both the equivalent notations $\phi(t, \omega, x)$ and $\phi(t, \omega)x$.

This definition of random cocycle follows the convention in e.g. [2,27]. In Arnold [2], the cocycle property
The skew-product structure $\Theta : \mathbb{R} \times \Omega \times X \to \Omega \times X$ characterizing the random dynamical system $(\theta, \phi)$ can be succinctly written as

$$\Theta(t)(\omega, x) := (\theta_t \omega, \phi(t, \omega, x)).$$

A probability measure $\mu$ on $(\Omega \times X, \mathcal{F} \otimes \mathcal{B}(X))$ is said to be invariant if

(i) $\mu(\Theta_t A) = \mu(A)$ for all $t \in \mathbb{R}$ and $A \in \mathcal{F} \otimes \mathcal{B}(X)$

(ii) The marginal of $\mu$ on $(\Omega, \mathcal{F})$ is $P$.

The canonical construction of the sample path space of Brownian motions can be briefly outlined as follows. Let $\Omega := C_0(\mathbb{R}, \mathbb{R}^d)$ be the space of all continuous functions $\xi : \mathbb{R} \to \mathbb{R}^d$ such that $\xi(0) = 0$, endowed with the compact open topology. Let $\mathcal{F} = \mathcal{B}(\Omega)$ denote the Borel $\sigma$-algebra on $\Omega$. Then, there exists the so-called Wiener probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ ensuring that the processes $(B(t))_{t \in \mathbb{R}}$ and $(W(t))_{t \in \mathbb{R}}$ are independent $d$-dimensional Brownian motions, with corresponding sample paths $(\omega, \beta) := \xi \in \Omega_{W} \times \Omega_{B}$ where $\Omega_{W}$ and $\Omega_{B}$ denote the intrinsic and common noise sample spaces. The natural filtration is the $\sigma$-algebra $\mathcal{F}_{s,t}$ generated by $\xi(u) - \xi(v)$ for $s \leq v \leq u \leq t$. The Wiener measure $\mathbb{P}$ is ergodic with respect to the Wiener shift map $\theta_t : \Omega \to \Omega$ defined by

$$(\theta_t \xi)(s) := \xi(s + t) - \xi(t), \quad s \in \mathbb{R}.$$ 

Therefore, $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t \in \mathbb{R}})$ is an ergodic random dynamical system.

With this sample path evolution as explicit representation of the noise, the SDE (1.1) is a random dynamical system on the product $\Omega \times \mathbb{R}^d$. We find that the stochastic Fokker-Planck equation (1.3) is a random dynamical system on $\Omega_B \times S$, where $S$ is the Schwartz space of solutions of (1.3) identified in Theorem 2.2.

**Proposition 4.1.** The stochastic Fokker-Planck equation (1.3) is a random dynamical system.

**Theorem 4.2 (Pullback attractor).** Let $\Phi$ be the random dynamical system associated to (1.3). Then, for $\mathbb{P}_B$-almost all $\beta \in \Omega_B$ (1.3) has a unique pullback attractor defined by

$$p_\beta := \lim_{\tau \to \infty} \Phi(\tau, \theta_{-\tau} \beta) p$$

which is independent of $p \in L^1$. Moreover, $p_\beta \in S$ and convergence is with respect to the semi-norm (2.5) on $S$.

By the Correspondence Theorem (see e.g. Arnold [2, Remark 1.4.2 and Proposition 1.4.3] and Crauel and Flandoli [17, Section 4]), if (1.1) has a unique stationary measure, then the

is required to hold for all $\omega \in \Omega$, instead of almost surely. In case the cocycle exists for almost all $\omega \in \Omega$ only, $\phi$ is called a crude cocycle and through a perfection procedure it possible to define an indistinguishable RDS for which the cocycle property is fulfilled for all noise realizations, see e.g. [21, Chapter 4.10] and references therein, most notably [28, 3].

15
corresponding random dynamical system has a unique invariant Markov measure, i.e. an invariant measure that is measurable with respect to the past \textsuperscript{12} and pullback attractors are identified with disintegrations of this Markov measure. We show in Proposition 4.4 that the common noise pullback attractors of (1.3) are equal to the expectation of the pullback attractors of (1.1) with respect to the intrinsic noise $W$, for a single fixed common noise realisation $\beta$. We summarize some well-established results on Markov measures and their disintegration \cite{2} in the context of our setting.

**Proposition 4.3 (Markov measure and its disintegration).** Let $\rho$ be the (unique) stationary measure of the random dynamical system $\varphi$ associated to (1.1) and $
abla \omega, \beta := \lim_{\tau \to \infty} \varphi(\tau, \theta_{-\tau} \omega, \theta_{-\tau} \beta) \ast \rho$. (4.1)

Then $\{\mu_{\omega, \beta}\}_{(\omega, \beta) \in \Omega}$ is $\mathbb{P}$-a.e. unique on $\mathcal{B}(\mathbb{R}^d)$ and

(i) for all $C \in \mathcal{B}(\mathbb{R}^d)$, $(\omega, \beta) \mapsto \mu_{\omega, \beta}(C)$ is $\mathcal{F}^-$-measurable, where $\mathcal{F}^- = \sigma(\bigcup_{s \leq t} \mathcal{F}_{s,t})$.

(ii) for $\mathbb{P}$-a.e. $(\omega, \beta) \in \Omega_W \times \Omega_B$, $\mu_{\omega, \beta}$ is a probability measure on $(X, \mathcal{B}(\mathbb{R}^d))$.

(iii) for all $A \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^d)$

$$
\mu(A) := \int_{\Omega_B} \int_{\Omega_W} \int_X 1_A(\omega, \beta, x) \mu_{\omega, \beta}(dx) \mathbb{P}_W(d\omega) \mathbb{P}_B(d\beta)
= \int_{\Omega_B} \int_{\Omega_W} \mu_{\omega, \beta}(A_{\omega, \beta}) \mathbb{P}_W(d\omega) \mathbb{P}_B(d\beta),
$$

where

$$A_{\omega, \beta} := \{x : (\omega, \beta, x) \in A\},$$

is an invariant probability measure of $\varphi$. The measure $\mu$ is known as the Markov measure associated to $\rho$ and it is the unique invariant probability measure of $\varphi$ that is measurable with respect to the past, cf. (i), such that

$$
\int_{\Omega_B} \int_{\Omega_W} \mu_{\omega, \beta} \mathbb{P}_W(d\omega) \mathbb{P}_B(d\beta) = \rho.
$$

Against this background, we now prove that the common noise pullback attractor of (1.3) is the expectation of the pullback attractor of the underlying SDE (1.1) with respect to the intrinsic noise.

**Proposition 4.4 (Common noise pullback attractor).** Let $\Phi$ be the random dynamical system associated to (1.3), $p_\rho$ be the Lebesgue density of the stationary measure $\rho$ of (1.1), and

$$
\mu_{\beta} := \int_{\Omega_W} \mu_{\omega, \beta} \mathbb{P}_W(d\omega).
$$

\textsuperscript{12}See Proposition 4.3(i).
Then, $\mathbb{P}_B$-a.s., $\mu_\beta$ is a probability measure on $\mathcal{B}(\mathbb{R}^d)$ with Lebesgue density $p_\beta$, where

$$p_\beta = \lim_{\tau \to \infty} \Phi(\tau, \theta_{-\tau} \beta)p_\rho$$

and

$$\int_{\Omega_B} p_\beta \mathbb{P}_B(d\beta) = p_\rho.$$

We finally stipulate a direct consequence of ergodicity for observables $g: \mathcal{S} \to \mathbb{R}$.

**Proposition 4.5.** Let $g: \mathcal{S} \to \mathbb{R}$ be continuous and integrable, then

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau g(\Phi(\tau, \beta)p)dt = \int_{\Omega_B} g(p_\beta)\mathbb{P}_B(d\beta),$$

(4.2)

$\mathbb{P}_B$-almost surely, independent of $p \in \mathcal{S}$.

**Appendices**

**A Proof of Theorem 2.2**

We start this section by demonstrating the equivalence between total mass conservation for measures and a weighted integrability condition to be fulfilled by a weak solution $q(t)$ of the initial value problem (2.2). As we shall discuss below, this condition will be employed to prove an important $L^1$ estimate (Lemma A.6) which, in turn, will be crucial for establishing uniqueness of solutions (Lemma A.7).

**Lemma A.1** (Equivalence between mass conservation and the weighted integrability condition). A weak solution $q$ of the non-autonomous Fokker-Planck equation (2.2) with initial condition $q_0 \in \mathcal{P}(\mathbb{R}^d)$ is a probability solution for any given $\beta \in \mathcal{C}^{1/2}(\mathbb{R}, \mathbb{R}^d)$, i.e.

$$q(t, x; \beta) \geq 0, \quad \int_{\mathbb{R}^d} q(t, x; \beta)dx = 1 \text{ for all } t \geq 0,$$

if and only if the weighted integrability condition

$$\lim_{N \to \infty} \int_0^T \int_{N < \|x\| < 2N} \|x\|^{-1} \|\nabla U\| \|q(t, x; \beta)\|dxdt = 0$$

(W.I.C.)

holds.

Proof. Let us define the test function $\varphi_N := \vartheta(\frac{x}{N})$ for any $N \in \mathbb{N}$, where $\vartheta \in C_c(\mathbb{R}^d, \mathbb{R}_+)$ is a cut-off function such that $\vartheta' \leq 0$ and

$$\vartheta(z) = \begin{cases} 
1 & \text{if } z \in [0, 1]^d \\
2 - z & \text{if } z \in [1, 2]^d.
\end{cases}$$
\(\vartheta\) is constructed to be at least \(C^2\) at \(z = 1\) and \(z = 2\). It is assumed to be equal to 0 for \(z \in [2, \infty)^d\) and extended evenly for \(z \in (-\infty, 0]^d\). Testing the non-autonomous Fokker Planck equation (2.2) with \(\varphi_N\), integrating by parts and rearranging terms we obtain

\[
\langle q(0), \varphi_N \rangle - \int_0^T \int_{\mathbb{R}^d} \nabla U \nabla \varphi_N q \, dx \, dt = \langle q(T), \varphi_N \rangle - \int_0^T \langle q, \Delta \varphi_N \rangle \, dt,
\]

where we suppressed the dependence on \((x, \beta)\) in order to simplify the notation. Since in the limit \(N \to \infty\) \(\Delta \varphi_N \sim N^{-2}\) and

\[
\lim_{N \to \infty} \langle q(T), \varphi_N \rangle = \|q(T)\|_{L^1},
\]

\[
\lim_{N \to \infty} \langle q(0), \varphi_N \rangle = \|q(0)\|_{L^1},
\]

we immediately obtain

\[
\lim_{N \to \infty} \int_0^T \int_{N \lVert x \rVert < 2N} \|x\|^{-1} \nabla U q \, dx \, dt = \|q(T)\|_{L^1} - \|q(0)\|_{L^1},
\]

and therefore we conclude.

\(\square\)

**Remark A.2** (Probability and bounded measures). The problems of uniqueness in the class of probability measures and in the class of all bounded measures are not equivalent. Consider for simplicity the one dimensional case, the autonomous scenario \(\beta \equiv 0\) and the potential

\[
V'(x) = 4x^3 + 16x^3(1 + 4x^4)^{-1}.
\]

Then, the Fokker-Planck equation (2.2) admits the following stationary, bounded sign-changing solution

\[
q(x) = x(1 + 4x^4)^{-1}.
\]

The weighted integrability condition is violated, since

\[
\int_{\mathbb{R}} x^2 |q(x)| \, dx = +\infty.
\]

At the same time, (2.2) admits the well known stationary probability solution

\[
\rho(x) = e^{-V(x)} = (1 + 4x^4)^{-1} e^{-x^4}.
\]

Therefore, there exists a unique solution in the class of probability measures, but there are also nonzero signed solutions in the class of bounded measures. For further details, see [13], Example 4.1.3.

Theorem 2.2 is proved by combining a series of energy-type estimates. We remark that, although the focus of this theorem is on probability measures, from here onwards we consider more broadly the evolution of signed measures. This is needed for the proof of uniqueness of the weak solution in Lemma A.7 where the evolution of the difference between two probability solutions is considered.

We make the following key assumptions:
Assumption A.3.

(I) Weak $L^1$ solutions of the Fokker Planck equation (2.2) are required to satisfy the weighted integrability condition (W.I.C.).

(II) The initial condition $q_0$ is a signed measure.

(III) The potential $U$ is infinitely differentiable and satisfies the dissipation condition (2.6).

Restricted to the setting of probability measures, Assumption A.3 boils down to the setting of Theorem 2.2.

Next, we note that the space $L^1(\mathbb{R}^d)$ can be interpreted as regular measures and embedded isometrically into the space of signed Borel measures $\mathcal{M}(\mathbb{R}^d)$. Although this result is already known, for the sake of having a self-contained discussion, we provide below an explicit proof, adapting the one in [42], Proposition 2.7.

Lemma A.4 (Approximation of signed measures). For every measure $\mu \in \mathcal{M}(\mathbb{R}^d)$, there exists a sequence of (signed) measures $(\mu_n)_{n \in \mathbb{N}} \in L^1(\mathbb{R}^d)$ such that

$$\|\mu_n\|_{L^1} \leq \|\mu\|_{\mathcal{M}(\mathbb{R}^d)}$$

for any $n \in \mathbb{N}$ and

$$\lim_{n \to \infty} \langle \mu_n, \varphi \rangle = \langle \mu, \varphi \rangle$$

for all test functions $\varphi \in C_0(\mathbb{R}^d)$, where $C_0(\mathbb{R}^d)$ denotes the space of all continuous functions with compact support on $\mathbb{R}^d$.

Proof. Let $(\varrho_n)_{n \in \mathbb{N}}$ be a sequence of mollifiers, that is, for every $n \in \mathbb{N}$, $\varrho_n \in C_\infty^0(\mathbb{R}^d)$, $\varrho_n$ is nonnegative, such that

$$\int_{\mathbb{R}^d} \varrho_n = 1$$

and for every $\delta > 0$

$$\lim_{n \to \infty} \int_{\mathbb{R}^d \setminus B(0,\delta)} \varrho_n = 0,$$

where $B(0,\delta)$ denotes the open ball centred at 0 with radius $\delta$. Then, we consider the convolution $\mu_n := \varrho_n * \mu$ between the mollifier and the measure $\mu \in \mathcal{M}(\mathbb{R}^d)$. We immediately have $\mu_n \in L^1(\mathbb{R}^d)$. From the convolution definition and thanks to Fubini’s theorem, we deduce

$$\int_{\mathbb{R}^d} \varphi d(\varrho_n * \mu) = \int_{\mathbb{R}^d} \varphi(x) \int_{\mathbb{R}^d} \varrho_n(x-y) d\mu(y) dx$$

$$= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varrho_n(x-y) \varphi(x) dx d\mu(y)$$

$$= \int_{\mathbb{R}^d} \varrho_n \varphi d\mu.$$
Applying again Fubini’s theorem we obtain
\[ \|\mu_n\|_{L^1} \leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \varrho_n(x - y)d\mu(y) \right) d\mu(x) \leq \int_{\mathbb{R}^d} d|\mu|(y) = \|\mu\|_{M(\mathbb{R}^d)} \]
and we conclude. \qed

Since our result will imply \( q(t) \in L^1(\mathbb{R}^d) \) for all \( t > 0 \), thanks to Lemma (A.4), in the proof of Theorem 2.2 we can restrict ourselves to the case \( q_0 \in L^1(\mathbb{R}^d) \).

Our strategy proceeds as follows. First, we prove the \( L^1 \) estimate (A.10) by employing the weighted integrability condition (W.I.C.). Thanks to (A.10), we establish uniqueness in Lemma (A.7). Next, we prove the weighted \( L^1 \)-estimate (A.12), which controls the behaviour of the tails of the solution. Following the same ideas, we prove the \( L^1 \)-localization estimate (A.15) and the two smoothing estimates (A.17) and (A.21). These results establish that the unique solution of the Fokker Planck equation (2.2) is smooth in the space variable \( x \), as regular in time as the function \( \beta \) and belongs to \( L^2(\mathbb{R}^d) \) at all times \( t > 0 \). Exploiting the structure of the equation and the infinite differentiability of the potential, we iterate the argument and achieve infinite differentiability and rapidly decreasing behaviour of all derivatives.

The \( L^1 \)-estimate (A.10) will be proved with the help of inequality (A.5) below. Given a weak solution \( q \) of the non-autonomous Fokker Planck equation (2.2), the fundamental idea to prove this inequality consists in regularizing the equation with a mollifier \( \vartheta_\delta \) first and then taking the limit \( \delta \to 0 \). Let \( \vartheta \in C^\infty_0(\mathbb{R}^d, \mathbb{R}_+) \) denote a non-negative mollification kernel satisfying \( \int_{\mathbb{R}^d} \vartheta(x)dx = 1 \) and define the standard Dirac’s delta approximation
\[ \vartheta_\delta(x) := \frac{1}{\delta} \vartheta\left(\frac{x}{\delta}\right). \]
Observe that, since \( \vartheta \) has a compact support,
\[ \nabla_x \vartheta_\delta(x - y) \neq 0 \quad \text{if} \quad |x - y| \leq C\delta \tag{A.1} \]
for some positive constant \( C \). Moreover, using integration by parts,
\[ \int_{\mathbb{R}^d} \vartheta'(s)sds = -\int_{\mathbb{R}^d} \vartheta(s)ds = -1. \tag{A.2} \]
We denote by \( S_\delta \) the mollification operator
\[ (S_\delta(q))(x) := \int_{\mathbb{R}^d} \vartheta_\delta(x - y)q(y)dy \tag{A.3} \]
and by \( \text{sgn}_\gamma(x) \) the standard smooth and monotone approximation of \( \text{sgn}(x) \), i.e.
\[ \text{sgn}_\gamma(x) := \frac{x}{\sqrt{x^2 + \gamma^2}}. \tag{A.4} \]
Finally, we define
\[ |z|_\gamma := \int_0^z \text{sgn}_\gamma(s)ds. \]
and notice that, by construction, \( \lim_{\gamma \to 0} |z|_\gamma = |z| \).
Lemma A.5 (Weak solution inequality). Any weak solution of (2.2) satisfies the inequality
\[
d\frac{dt}{dt} \| q(\cdot, t; \beta) \|, \varphi \leq (\partial_t \varphi - \nabla U \nabla \varphi + \Delta \varphi, \| q(\cdot, t; \beta) \|) \quad (A.5)
\]
for almost all \( t \geq 0 \), any \( \beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d) \) and \( \varphi \in C^\infty_0((0, T) \times \mathbb{R}^d) \) such that \( \varphi \geq 0 \).

Proof. We apply the mollification operator \( S_\delta \) defined in (A.3) to both sides of the non-autonomous Fokker-Planck equation (2.2), thereby obtaining
\[
\partial_t (S_\delta(q)) = \nabla(S_\delta(\nabla U q)) + \Delta(S_\delta(q)), \quad S_\delta(q)|_{t=0} = S_\delta(q_0).
\]
We define the test function
\[
\psi(t, x) := \varphi(t, x) \text{sgn}_\gamma(S_\delta(q)(x)),
\]
where \( \varphi \) is the test function in (A.5) and \( \text{sgn}_\gamma \) is the smooth approximation of the \( \text{sgn} \) function defined in (A.4). We obtain
\[
\frac{d}{dt} \langle |S_\delta(q)|_\gamma, \psi \rangle = \langle \partial_t \psi, |S_\delta(q)|_\gamma \rangle + \langle \nabla(S_\delta(\nabla U q)), \psi \rangle + \langle \Delta S_\delta(q), \psi \rangle, \quad (A.6)
\]
The idea now is to write each term on the RHS of (A.6) in a convenient form by means of integration by parts and take the limit \( \delta \to 0 \). Firstly, integration by parts implies term (a) can be rewritten as
\[
\langle \Delta S_\delta(q), \psi \rangle = -\langle \nabla S_\delta(q), \text{sgn}'_\gamma(S_\delta(q)) \nabla S_\delta(q) \varphi \rangle - \langle \nabla S_\delta(q), \text{sgn}_\gamma(S_\delta(q)) \nabla \varphi \rangle, \quad (A.7)
\]
where \( ' \) denotes the derivative, since boundary terms vanish thanks to the dissipation condition (2.6). Dropping the first term on the RHS of (A.7) and applying integration by parts again yields
\[
\langle S_\delta(q), \psi \rangle \leq \langle \Delta \varphi, |S_\delta(q)|_\gamma \rangle.
\]
Taking the limit,
\[
\lim_{(\gamma, \delta) \to (0, 0)} \langle \Delta \varphi, |S_\delta(q)|_\gamma \rangle = \langle |q|, \Delta \varphi \rangle.
\]
Next, term (b) in (A.6) reads as
\[
\langle \nabla(S_\delta(\nabla U q)), \varphi \text{sgn}_\gamma(S_\delta(q)) \rangle = \langle \nabla(\nabla U S_\delta(q)), \varphi \text{sgn}_\gamma(S_\delta(q)) \rangle \quad (A.8a)
\]
\[
+ \langle \nabla(S_\delta(\nabla U q) - \nabla U S_\delta(q)), \varphi \text{sgn}_\gamma(S_\delta(q)) \rangle. \quad (A.8b)
\]
The first term (A.8a) can be written as
\[
\langle \nabla(\nabla U S_\delta(q)), \varphi \text{sgn}_\gamma(S_\delta(q)) \rangle = \langle \Delta U \varphi, S_\delta(q) \text{sgn}_\gamma(S_\delta(q)) \rangle + \langle \nabla U \psi, \nabla |S_\delta(q)|_\gamma \rangle
\]
\[
= \langle \Delta U \varphi, S_\delta(q) \text{sgn}_\gamma(S_\delta(q)) \rangle - \langle \nabla(\nabla U \varphi), |S_\delta(q)|_\gamma \rangle
\]
\[
\to \langle \Delta U \varphi, |q| \rangle - \langle \nabla(\nabla U \varphi), |q| \rangle
\]
\[
= -\langle \nabla \nabla U \varphi, |q| \rangle
\]

21
as \((\gamma, \delta) \to (0, 0)\). In order to write more explicitly the term \((A.8b)\), we recall that
\[
S_\delta(\nabla Uq)(x) - \nabla US_\delta(q)(x) = \int_{\mathbb{R}^d} \vartheta_\delta(x-y) [\nabla_y U - \nabla x U] q(y)dy
\]
and therefore
\[
\langle \nabla (S_\delta(\nabla Uq) - \nabla US_\delta(q)), \varphi \rangle_{L^2} = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \vartheta_\delta(x-y) (\nabla_y U - \nabla x U) q(y) \varphi(x) \text{sgn}_\gamma(S_\delta(q)(x)) dy dx.
\]
Taking the limit \(\delta \to 0\) in \((A.9b)\) yields
\[
\lim_{\delta \to 0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \vartheta_\delta(x-y) \nabla U(x) q(y) \varphi(x) \text{sgn}_\gamma(S_\delta(q)(x)) dy dx = \langle \nabla U \varphi, \text{sgn}_\gamma(q) \rangle.
\]
For what concerns the integral term \((A.9a)\), instead, we use \((A.1)\). Consequently, we may write
\[
\nabla_y U - \nabla x U = -\Delta U(x)(x-y) + O(\delta^2(\|x\| + 1)).
\]
Thus, at first order approximation, \((A.9a)\) reduces to
\[
- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \vartheta_\delta(x-y) \Delta U(x)(x-y) q(y) \varphi(x) \text{sgn}_\gamma(S_\delta(q)(x)) dy dx.
\]
Thanks to \((A.2)\), we deduce
\[
\lim_{\delta \to 0} - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \vartheta_\delta(x-y) \Delta U(x)(x-y) q(y) \varphi(x) \text{sgn}_\gamma(S_\delta(q)(x)) dy dx = \langle \Delta U \varphi, \text{sgn}_\gamma(q) \rangle.
\]
Hence, \((A.8b)\) vanishes as \(\delta \to 0\). Putting everything together, we let \((\gamma, \delta) \to (0, 0)\), observe that \(\psi \to \varphi\) and \(|S_\delta(q)|_\gamma \to |q|\) and finally obtain the inequality \((A.5)\). \(\square\)

We are now ready to prove the following \(L^1\)-estimate, the proof of which heavily relies on the weighted integrability condition \((W.I.C.)\).

**Lemma A.6** \((L^1\)-estimate\). *Any weak solution of the non-autonomous Fokker Planck equation \((2.2)\) satisfies for any given \(\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)\) the \(L^1\)-estimate*
\[
||q(\cdot, t; \beta)||_{L^1} \leq ||q(\cdot, 0; \beta)||_{L^1}, \quad t > 0.
\]

**Proof.** Let us consider a cut-off function \(\vartheta \in C^\infty_0(\mathbb{R}^d, [0, 1])\) such that
\[
\vartheta(z) = \begin{cases} 
1 & \text{if } z \in [-1, 1]^d \\
0 & \text{if } z \in [2, \infty)^d 
\end{cases}
\]
and the test function $\varphi_N := \vartheta \left( \frac{x}{N} \right)$ for some $N \in \mathbb{N}$. Then, using the fact that the function $t \to \langle |q(t)|, \varphi_N \rangle$ is continuous in time, inequality (A.5) implies

$$
\langle |q(T)|, \varphi_N \rangle \leq \langle |q(\tau)|, \varphi_N \rangle + \int_\tau^T \langle \Delta \varphi_N - \nabla U \nabla \varphi_N, \|q\| \rangle dt,
$$
(A.11)

where we suppressed the dependence on $x$ and $\beta$ in order to simplify the notation. The weighted integrability condition (W.I.C.) ensures that the integral on the RHS of (A.11) tends to 0 as $N \to +\infty$. Indeed,

$$
\left| \langle \nabla U \nabla \varphi_N, \|q\| \rangle \right| \leq \frac{1}{N} \int_{N < \|x\| < 2N} \|\nabla U\| \left| q' \left( \frac{x}{N} \right) \right| \|q\| dx
\leq C \int_{N < \|x\| < 2N} \|x\|^{-1} \|\nabla U\| \|q\| dx
$$

and

$$
\int_\tau^T \langle \nabla U \nabla \varphi_N, \|q\| \rangle dt \leq \int_\tau^T \int_{N < \|x\| < 2N} \|x\|^{-1} \|\nabla U\| \|q\| dx dt \to 0,
$$
as $N \to \infty$. Passing to this limit in (A.11) therefore yields

$$
\|q(\tau)\|_{L^1} \leq \|q(0)\|_{L^1}.
$$

Finally, we let $\tau \to 0$, exploit continuity and conclude.

Uniqueness of the weak solution for the initial value problem (2.2) is an immediate consequence of Lemma A.6.

**Lemma A.7 (Uniqueness).** The initial value problem (2.2) admits a unique weak solution for any given $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$.

**Proof.** Let us denote by $q_1, q_2$ two distinct weak solutions of (2.2) with the same initial condition $q_0$. Define $q(x, t; \beta) := q_1(x, t; \beta) - q_2(x, t; \beta)$. Then, thanks to linearity, $q$ will be a weak solution of (2.2) and

$$
\|q(\cdot, t; \beta)\|_{L^1} \leq \|q(\cdot, 0; \beta)\|_{L^1} = 0
$$

thanks to the $L^1$-estimate (A.10).

Next, we proceed with proving the following weighted $L^1$-estimate, which describes the global behaviour of the tails of the weak solution.

**Lemma A.8 (Weighted $L^1$-estimate).** The unique weak solution of the non-autonomous Fokker Planck equation (2.2) satisfies for any given $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$ the weighted $L^1$-estimate

$$
\| (1 + x^n) q(\cdot, t; \beta) \|_{L^1} \leq C(\beta) \| (1 + x^n) q(\cdot, 0; \beta) \|_{L^1},
$$
(A.12)

for any $n \in \mathbb{N}$ and some constant $C = C(\beta) < \infty$.\(^{13}\)

\(^{13}\)In this and subsequent lemmas time-dependence of the constant $C$ is not problem. Since we are interested in local regularity for $t > 0$, we are considering $t \in [0, T]$. Without loss of generality, we might set $T = 1$. For what concerns the $\beta$-dependence instead, the constant $C$ in general will not be uniform with respect to $\beta$. To gain uniformity, additional assumptions on $\beta$ would be required (such as boundedness), but in our context this is not needed.
Proof. We multiply the non-autonomous Fokker Planck equation \((2.2)\) by \((1 + x^n)\text{sgn}(q)\) for any \(n \in \mathbb{N}\) and integrate over \(\mathbb{R}^d\):

\[
\frac{d}{dt} \|(1 + x^n)q\|_{L^1} = \int_{\mathbb{R}^d} (1 + x^n)\text{sgn}(q) \partial_t q dx
\]

\[= \int_{\mathbb{R}^d} (1 + x^n)\text{sgn}(q) \Delta q dx + \int_{\mathbb{R}^d} (1 + x^n)\text{sgn}(q) \nabla(\nabla U q) dx.\]

Integration by parts applied to the term \((b)\) yields

\[
\int_{\mathbb{R}^d} (1 + x^n)\text{sgn}(q) \nabla(\nabla U q) dx = \nabla U q(1 + x^n)\text{sgn}(q) \bigg|^{+\infty}_{-\infty} - \int_{\mathbb{R}^d} \nabla U q \text{sgn}(q) nx^{n-1} dx
\]

\[\leq -\gamma \|(1 + x^{3+(n-1)})q\|_{L^1} + C(\beta)\|q\|_{L^1},\]

for some constants \(\gamma > 0\) and \(C(\beta) > 0\) depending on the modulus of \(\beta\), where we used the dissipation condition for the shifted potential \((2.6)\). Similarly, integration by parts applied to the term \((a)\) yields

\[
\int_{\mathbb{R}^d} (1 + x^n)\text{sgn}(q) \Delta q dx = -\int_{\mathbb{R}^d} (\nabla q)\text{sgn}(q) nx^{n-1} dx
\]

\[\leq -n \int_{\mathbb{R}^d} \nabla \|q\| x^{n-1} dx
\]

\[= n(n-1) \int_{\mathbb{R}^d} \|q\| x^{n-2} dx = n(n-1)\|x^{n-2} q\|_{L^1}.\]

Hence,

\[
\frac{d}{dt} \|(1 + x^n)q\|_{L^1} + \gamma \|(1 + x^{3+(n-1)})q\|_{L^1} \leq C(\beta)\|q\|_{L^1}
\]

Setting \(\gamma = 0\), integrating with respect to time, using the \(L^1\)-estimate \((A.10)\) and noticing that

\[
\|(1 + x^n)q(\cdot, 0; \beta)\|_{L^1} \geq \|q(\cdot, 0; \beta)\|_{L^1},
\]

we finally obtain inequality \((A.12)\).

Next, given a weighted weak solution of our non-autonomous Fokker-Planck equation, we establish a \(L^1\)-localization estimate.

**Lemma A.9 (\(L^1\) localization estimate).** The unique weak solution of the non-autonomous Fokker Planck equation \((2.2)\) satisfies for any given \(\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)\) the \(L^1\) localization estimate

\[
\|(1 + x^n)q(\cdot, t; \beta)\|_{L^1} \leq C(\beta) \frac{1 + t^N}{t^N} \|q(\cdot, 0; \beta)\|_{L^1},
\]

for any \(n \in \mathbb{N}, t > 0, \) some \(N \in \mathbb{N}\) and some constant \(C = C(\beta) < \infty\).

**Proof.** Let us consider again the inequality

\[
\frac{d}{dt} \|(1 + x^n)q\|_{L^1} + \gamma \|(1 + x^{3+(n-1)})q\|_{L^1} \leq C(\beta)\|q\|_{L^1}
\]
for any $n \in \mathbb{N}$ and some constant $\gamma > 0$, as derived in the proof of Lemma A.8. We multiply both sides by $t^N$ for some $N \in \mathbb{N}$:

$$t^N \frac{d}{dt}(1 + x^n)q \|_{L^1} + t^N \gamma(1 + x^{3+(n-1)})q \|_{L^1} \leq t^N C(\beta)q \|_{L^1}.$$ 

Using the Hölder inequality

$$t^{N-1}(1 + x^n)q \|_{L^1} \leq \|q\|_{L^1} \sqrt{\frac{N}{t^N} (t^N \sqrt{\frac{N}{t^N}} q \|_{L^1})^{\frac{N-1}{N}}} ,$$  \hspace{1cm} (A.16)

we obtain

$$\frac{d}{dt}(t^N(1 + x^n)q \|_{L^1}) + \gamma t^N(1 + x^{3+(n-1)})q \|_{L^1} \leq C(\beta)(1 + t^N)q \|_{L^1},$$

which implies

$$\frac{d}{dt}(t^N(1 + x^n)q \|_{L^1}) \leq C(\beta)(1 + t^N)q \|_{L^1} \leq C(\beta)(1 + t^N)q_0 \|_{L^1},$$

thanks to the $L^1$-estimate (A.10). Integrating on both sides, we conclude. \hfill \Box

Next, we prove two smoothing estimates which ensure the unique weak solution of the initial value problem (2.2) belongs to $L^2$ and, in fact, to the Sobolev space $H^1 := W^{1,2}$, for all $t > 0$, that is, its spatial derivative belongs to $L^2$ as well.

**Lemma A.10 (First smoothing estimate).** The unique weak solution of the non-autonomous Fokker Planck equation (2.2) satisfies for any given $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$

$$\|\cdot(t; \beta)\|_{L^2}^2 + \int_t^{t+1} \|\nabla s q(s; \beta)\|_{L^2}^2 ds \leq C(\beta) \frac{t^{N+1}}{t^N} \|\cdot(0; \beta)\|_{L^1}^2 ,$$  \hspace{1cm} (A.17)

for some $N \in \mathbb{N}$ and some constant $C = C(\beta) < \infty$.

**Proof.** Let us multiply the non-autonomous Fokker Planck equation (2.2) by $q$ and integrate over $\mathbb{R}^d$:

$$\int_{\mathbb{R}^d} q \partial_t q dx = \int_{\mathbb{R}^d} q \Delta q dx + \int_{\mathbb{R}^d} q \nabla(\nabla U) dx .$$

Using integration by parts we obtain

$$\int_{\mathbb{R}^d} q \Delta q dx = \nabla q \nabla q \bigg|^{+\infty}_{-\infty} - \int_{\mathbb{R}^d} (\nabla q)^2 dx = -\|\nabla q\|_{L^2}^2$$

$$\int_{\mathbb{R}^d} q \nabla(\nabla U) dx = \frac{1}{2} \int_{\mathbb{R}^d} \|q\|_{L^2}^2 \Delta U dx .$$

Putting everything together,

$$\frac{1}{2} \frac{d}{dt} \|q\|_{L^2}^2 + \|\nabla q\|_{L^2}^2 = \frac{1}{2} \langle \Delta U, \|q\|_{L^2}^2 \rangle .$$

25
Thanks to the dissipation condition (2.6) on the potential $U$, we obtain
\[
\frac{1}{2} \langle \Delta U, q^2 \rangle \leq C(\beta) (1 + \|x\|^2, q^2) \leq C(\beta) \|(1 + \|x\|^2)q\|_{L^1} \|q\|_{L^\infty}.
\]

Next, we use the inequality
\[
\|q\|_{L^\infty}^2 \leq \|q\|_{L^2} \|\nabla q\|_{L^2}
\]
and deduce
\[
\frac{1}{2} \frac{d}{dt} \|q\|_{L^2}^2 + \|\nabla q\|_{L^2}^2 \leq C(\beta) \|(1 + \|x\|^2)q\|_{L^1} \|q\|_{L^\infty}
\]
\[
\leq C(\beta) \|(1 + \|x\|^2)q\|_{L^1} \|q\|_{L^2}^{1/2} \|\nabla q\|_{L^2}^{1/2}.
\]

Using also $ab \leq \frac{1}{2}(\epsilon^2 a^2 + \epsilon^{-2} b^2)$ for $\epsilon$ small enough, we obtain
\[
\frac{d}{dt} \|q\|_{L^2}^2 + \|\nabla q\|_{L^2}^2 \leq C(\beta) \|q\|_{L^2}^2 + C(\beta) \|(1 + \|x\|^2)q\|_{L^1}^2.
\]

Using inequality (A.19), we also have
\[
\|q\|_{L^2}^2 \leq \|q\|_{L^1} \|q\|_{L^\infty}
\]
\[
\leq \|q\|_{L^1} \|q\|_{L^2}^{1/2} \|\nabla q\|_{L^2}^{1/2},
\]
from which we deduce
\[
\|q\|_{L^2} \leq \|q\|_{L^1}^{2/3} \|\nabla q\|_{L^1}^{1/3}.
\]

Putting everything together, we obtain the relationship
\[
\frac{d}{dt} \|q\|_{L^2}^2 + \|\nabla q\|_{L^2}^2 \leq C(\beta) \|(1 + x^2)q\|_{L^1}.
\]

Multiplying by $t^N$ for some $N \in \mathbb{N}$, using again inequality (A.20) and integrating with respect to time from $t$ to $t + 1$, we conclude.

\[\text{Lemma A.11 (Second smoothing estimate).} \quad \text{The unique weak solution of the non-autonomous Fokker Planck equation} \ (2.2) \ \text{satisfies for any given} \ \beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)\]
\[
\|q(\cdot, t; \beta)\|_{H^1} \leq C(\beta)^{1 + \frac{1}{t^{N}}} \|q(\cdot, 0; \beta)\|_{L^1},
\]
\[\text{for some} \ N \in \mathbb{N} \ \text{and some constant} \ C = C(\beta) < \infty.\]

\[\text{Proof.} \quad \text{Let us multiply the non-autonomous Fokker Planck equation} \ (2.2) \ \text{by} \ \Delta q \ \text{and integrate over} \ \mathbb{R}^d:\]
\[
\int_{\mathbb{R}^d} \partial_t q \Delta q dx = \int_{\mathbb{R}^d} (\Delta q)^2 dx + \int_{\mathbb{R}^d} \Delta q \nabla (\nabla U q) dx. \tag{A.22}
\]

Terms (a) and (b) in (A.22) can be rewritten respectively as
\[
\int_{\mathbb{R}^d} \partial_t q \Delta q dx = -\frac{1}{2} \frac{d}{dt} \|\nabla q\|_{L^2}^2
\]
\[
\int_{\mathbb{R}^d} (\Delta q)^2 dx = \|\Delta q\|_{L^2}^2.
\]
Regarding term (c), using again integration by parts and vanishing at the boundary due to the dissipation condition, we deduce
\[
\int_{\mathbb{R}^d} \Delta q \nabla (\nabla U q) dx = - \int_{\mathbb{R}^d} \nabla q \nabla (\nabla U q) dx \\
= - \int_{\mathbb{R}^d} \nabla q \left( \Delta U q + 2 U \nabla q + \nabla U \Delta q \right) dx
\]
\[
= - \int_{\mathbb{R}^d} \nabla q \Delta U q dx - 2 \int_{\mathbb{R}^d} \nabla q \Delta U dx - \int_{\mathbb{R}^d} \nabla q \nabla U \Delta q dx.
\]
Using
\[
\int_{\mathbb{R}^d} \Delta U \nabla \nabla q dx = - \int_{\mathbb{R}^d} q \Delta U q dx - \int_{\mathbb{R}^d} \Delta q U dx,
\]
we obtain
\[
\int_{\mathbb{R}^d} \Delta q \nabla (\nabla U q) dx = - \int_{\mathbb{R}^d} \Delta U \nabla \nabla q dx + \int_{\mathbb{R}^d} \Delta q U dx - \int_{\mathbb{R}^d} \nabla q \nabla U \Delta q dx.
\]
We further notice that
\[
\int_{\mathbb{R}^d} \nabla q \Delta q \nabla U dx = - \frac{1}{2} \int_{\mathbb{R}^d} \nabla q \nabla U dx.
\]
Putting everything together, we find
\[
\int_{\mathbb{R}^d} \nabla (\nabla U q) \Delta q dx = - \frac{1}{2} \int_{\mathbb{R}^d} \Delta U \nabla \nabla q dx + \int_{\mathbb{R}^d} \Delta q U dx \\
= \frac{3}{2} \int_{\mathbb{R}^d} \Delta q U dx + \frac{1}{2} \int_{\mathbb{R}^d} \nabla^3 U \nabla q dx.
\]
Concisely, the equality above can be written as
\[
\langle \nabla (\nabla U q), \Delta q \rangle_{L^2} = \frac{3}{2} \langle \Delta q, \Delta U q \rangle_{L^2} + \frac{1}{2} \langle \nabla^3 U \nabla q, q \rangle_{L^2}.
\]
Therefore, using the inequality
\[
\| \nabla q \|^2_{L^2} \leq \| \Delta q \|_{L^2} \| q \|^2_{L^2}
\]
we deduce
\[
\| \nabla (\nabla U q), \Delta q \|_{L^2} \leq \epsilon \| \Delta q \|^2_{L^2} + C(\beta) \| \nabla q \|^2_{L^2} + C(\beta) \| (1 + x^2) q \|^2_{L^2} \\
\leq \epsilon \| \Delta q \|^2_{L^2} + C(\beta) \| \nabla q \|^2_{L^2} + C(\beta) \| (1 + x^2) q \|_{L^1} \| q \|_{L^\infty} \\
\leq \epsilon \| \Delta q \|^2_{L^2} + C(\beta) \left( \| \nabla q \|^2_{L^2} + \| q \|^2_{L^2} + \| (1 + x^4) q \|_{L^1} \| q \|^2_{L^\infty} \right) \\
\leq 2\epsilon \| \Delta q \|^2_{L^2} + C(\beta) \left( \| \nabla q \|^2_{L^2} + \| q \|^2_{L^2} + \| (1 + x^4) q \|_{L^1} \| q \|^2_{L^\infty} \right)
\]
for \( \epsilon > 0 \) small enough and some positive constant \( C = C(\beta) \). This implies
\[
\frac{d}{dt} \| \nabla q \|^2_{L^2} + \gamma \| \Delta q \|^2_{L^2} + \| \nabla q \|^2_{L^2} \leq C(\beta) \left( \| \nabla q \|^2_{L^2} + \| q \|^2_{L^2} + \| (1 + x^4) q \|_{L^1} \| q \|_{L^\infty} \right),
\]
for some constant \( \gamma > 0 \). Multiplying by \( t^N \) for some \( N \in \mathbb{N} \), using again inequality \( \text{(A.26)} \), integrating with respect to time and, finally, employing the first smoothing estimate \( \text{(A.17)} \), we conclude. \( \square \)
Finally, we are ready to prove the main theorem of this section. Having established existence, uniqueness and all estimates above for signed measures, we now restrict ourselves to probability measures. We denote by $H^k := W^{k,2}$ the Hilbert space of all functions $f \in L^2$ such that their weak derivatives up to order $k$ have finite $L^2$ norm. We further denote by $L^1_{1+x^n}$, for any $n \in \mathbb{N}$, the weighted space of measurable functions $f$ such that

$$\|f\|_{L^1_{1+x^n}} := \int_{\mathbb{R}^d} (1 + x^n) \|f(x)\| dx < \infty.$$ 

**Proof of Theorem 2.2.** Combining the $L^1$ estimates (A.10), (A.12), (A.15) and the smoothing estimates (A.17), (A.21), we conclude that for any given $\beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$, $q(t) \in L^1_{1+x^n} \cap H^1$ for any $n \in \mathbb{N}$ and $t > 0$. In order to gain more regularity, we differentiate the Fokker-Planck equation (2.2) with respect to the space variable $x$ iteratively, via a standard bootstrapping procedure. After one step, we obtain $\partial_x q(t) \in L^1_{1+x^n} \cap H^1$, which implies in particular $q(t) \in H^2$ for $t > 0$. After two steps, we obtain $\partial_x^2 q(t) \in L^1_{1+x^n} \cap H^1$, which implies $q(t) \in H^3$ for $t > 0$ and so on. This shows that, at any time $t > 0$, $q(t) \in H^\infty$ and any spatial derivative $\partial_x^\alpha q(t)$ decays faster than any polynomial as $|x| \to \infty$, at any time $t > 0$. Finally, we recall that the Sobolev space $H^s(\mathbb{R}^d)$ can be continuously embedded into $C^k(\mathbb{R}^d)$ for any $k \in \mathbb{N}$ and $s > k + n/2$ [15]. This readily implies $H^\infty$ can be continuously embedded in $C^\infty$. Putting everything together, we conclude that $q(t)$ belongs to the Schwarz space $S$ for any time $t > 0$. □

**B Proofs of Propositions 3.2 and 3.3**

**Lemma B.1** (Dissipation condition and strict convexity). Assume the potential $V$ satisfies the dissipation condition (2.6), that is,

$$\nabla V(x) \cdot x \|x\|^2 \geq \frac{1}{2} \|x\|^6 - C$$

for some $C > 0$. Then $V$ is strictly convex outside a given ball in $\mathbb{R}^d$.

**Proof.** We rewrite the dissipation condition as

$$\nabla V(x) = Mx\|x\|^2 + h(x),$$

with $M \geq \frac{1}{2}, h(x) = O(\|x\|^2)$ and such that

$$f(x) := \left(M - \frac{1}{2}\right) \|x\|^6 + h(x) \cdot x \|x\|^2$$

is lower bounded. Differentiating (B.1) yields

$$\Delta V(x) = 3M\|x\|^2 + g(x),$$

with $g(x) = \nabla \cdot h(x) = O(\|x\|)$, meaning that there exists $L > 0$ and $x_0 \in \mathbb{R}$ such that for any $\|x\| \geq x_0$

$$|g(x)| \leq L\|x\|.$$
Then, letting \( N := 3M \), we have
\[
\Delta V(x) = N\|x\|^2 + g(x) \geq N\|x\|^2 - L\|x\|
\]
and so the potential \( V \) is strictly convex outside the ball centered in \( \frac{LN}{2} \) with radius \( \frac{LN}{2} \).

We illustrate this result in the context of the example of Section [1.3.2].

**Example 1** (Double well potential). We consider the one-dimensional double-well potential 
\( V'(x) = x(x^2 - a) \), with \( a > 0 \). Then, the dissipation condition \( \text{(2.6)} \) is fulfilled if and only if
\[
x^4 \left( \frac{1}{2} x^2 - a \right) \geq -C
\]
for some constant \( C > 0 \). Let \( f(x) := \frac{1}{2}x^6 - x^4a \). Then, we set
\[
f'(x) = 3x^5 - 4ax^3 = x^3(3x^2 - 4a) = 0
\]
and we find the local extrema \( x = 0 \) and \( x = \pm \sqrt{\frac{4a}{3}} \). Hence, \( f(0) = 0 \) and
\[
f \left( \pm \sqrt{\frac{4a}{3}} \right) = a^3 \left( -\frac{16}{27} \right) \geq -C
\]
for any constant \( C \geq \frac{16}{27}a^3 \). The dissipation condition is therefore satisfied. Moreover,
\[
V''(x) = 3x^2 - a > 0 \iff x^2 > \frac{a}{3}
\]
and so we immediately deduce the potential \( V \) is strictly convex outside the ball centred at 0 with radius \( \sqrt{\frac{a}{3}} \).

The converse implication of Lemma \((\text{B.1})\) does not hold, as the following example shows.

**Example 2** (Strict convexity outside a ball does not imply the dissipation condition). We consider the one-dimensional potential
\[
V(x) = \frac{x^4}{8} - a\frac{x^2}{2}, \quad a > 0.
\]
Then,
\[
V'(x) = x^3 - ax = x \left( \frac{x^2}{2} - a \right)
\]
\[
V''(x) = \frac{3}{2}x^2 - a > 0 \iff x^2 > \frac{2a}{3}.
\]
We immediately see that \( V \) is strictly convex outside the ball centred at 0 with radius \( \sqrt{\frac{2a}{3}} \). However, the dissipation condition reads as
\[
V'(x)x^3 = \frac{1}{2}x^6 - ax^4 \geq \frac{1}{2}x^6 - C
\]
which is equivalent to \( C \geq ax^4 \) for some positive constant \( C \) and this is clearly not possible.
Proof of Proposition 3.2. We extend results by Eberle [22] to a non-autonomous setting. Consider the difference process \( z(t) := x(t) - y(t) \). Then,

\[
dz(t) = (\nabla V(y) - \nabla V(x))dt + 2|\sigma^{-1}z(t)|^{-1}z(t)d\tilde{W}(t), \quad t < T
\]

and \( z(t) = 0 \) for \( t \geq T \), where

\[
\tilde{W}(t) := \int_0^t e^T(s)dW(s)
\]
is a new Brownian motion by Levy’s characterization, \( e(t) \) is the unit vector defined by

\[
e(t) := \frac{\sigma^{-1}(x(t) - y(t))}{|\sigma^{-1}(x(t) - y(t))|}
\]

and \( T \) is the coupling time. Next, we define \( r(t) := \|z(t)\| = |\sigma^{-1}z(t)| \). By application of Ito’s formula we find

\[
dr(t) = 2|\sigma^{-1}z(t)|^{-1}r(t)d\tilde{W}(t) + r^{-1}(t)z(t)(\sigma\sigma^T)^{-1}(\nabla V(x(t)) + \nabla V(y(t)))dt.
\]

Given a smooth function \( f \in C^1(\mathbb{R}^d) \), this implies

\[
df(r(t)) = 2|\sigma^{-1}z(t)|^{-1}r(t)f'(r(t))d\tilde{W}(t)
+ r^{-1}(t)z(t)(\sigma\sigma^T)^{-1}(\nabla V(x(t)) + \nabla V(y(t)))f'(r(t))dt
+ 2|\sigma^{-1}z(t)|^{-2}r^2(t)f''(r(t))dt. \tag{B.2}
\]

We also define for any \( r > 0 \) the function

\[
k(r) := \inf_{x, y \in \mathbb{R}^d, \|x - y\| = r} \left\{ -2|\sigma^{-1}(x - y)|^2 \frac{(x - y) \cdot (\sigma\sigma^T)^{-1}(\nabla V(x) + \nabla V(y))}{\|x - y\|^2} \right\} \tag{B.3}
\]

Indeed, \( k(r) \) is the largest positive real number such that

\[
(x - y) \cdot (\sigma\sigma^T)^{-1}(\nabla V(x) + \nabla V(y)) \leq -\frac{1}{2}k(r)\|x - y\|^2
\]

for any \( x, y \in \mathbb{R}^d \) with \( \|x - y\| = r \). Let us denote by \( m(t) \) the drift on the right hand side of (B.2). By definition of \( k \),

\[
m(t) \leq \Gamma(t) := 2|\sigma^{-1}z(t)|^{-2}r^2(t) \cdot \left( f''(r(t)) - \frac{1}{4}k(r(t))f'(r(t)) \right).
\]

Hence, the process \( e^{ctf(r(t))} \) is a supermartingale for \( t < T \) if \( \Gamma(t) \leq -cf(r(t)) \). We aim to find a constant \( c \) and function \( f \) such that this inequality holds. Let

\[
\alpha := \sup \left\{ |\sigma^{-1}z|^2 : z \in \mathbb{R}^d \text{ with } \|z\| = 1 \right\}.
\]
Since for any $z \in \mathbb{R}^d$

$$|\sigma^{-1}z|^2 \leq \alpha \|z\|^2,$$

it suffices for $f$ to satisfy

$$f''(r) - \frac{1}{4}rk(r)f'(r) \leq -\frac{\alpha c}{2} f(r) \quad \text{(B.4)}$$

for all $r > 0$, cf. [22, eq. 63]. We observe this equation holds with $c = 0$ in case

$$f'(r) = \varphi(r) := \exp\left(-\frac{1}{4} \int_0^r sk^-(s)ds\right),$$

where $k^- := \max\{-k, 0\}$ denotes the negative part of the function $k$. Next, following [22], we make the ansatz

$$f'(r) = \varphi(r)g(r), \quad \text{(B.5)}$$

where $g \geq \frac{1}{2}$ is a decreasing absolutely continuous function satisfying $g(0) = 1$. Notice that the condition $g \geq 0$ is necessary to ensure that $f$ is non-decreasing. The condition

$$\frac{1}{2} \geq g \geq 1$$

ensures

$$\frac{\Phi}{2} \leq f \leq \Phi, \quad \Phi(r) := \int_0^r \varphi(s)ds.$$

The ansatz (B.5) yields

$$f''(r) = -\frac{1}{4}k^-(r)f(r) + \varphi(r)g(r) \leq \frac{1}{4}rk(r)f(r) + \varphi(r)g'(r).$$

In turn, condition (B.4) is satisfied if

$$g'(r) \leq -\frac{\alpha c f(r)}{2 \varphi(r)}. \quad \text{(B.6)}$$

Next, we define two constants $R_0, R_1 \geq 0$, with $R_0 \leq R_1$:

$$R_0 := \inf \{R \geq 0 : k(r) \geq 0, \forall r \geq R\}$$
$$R_1 := \inf \{R \geq R_0 : k(r)R(R - R_0) \geq 8, \forall r \geq R\}.$$

As remarked in [22], we can rewrite $k$ as

$$k(r) = \inf \left\{2 \int_0^1 \partial_{(x-y)/|x-y|}(\sigma\sigma^\top)^{-1}V((1-t)x + ty)dt : x, y \in \mathbb{R}^d s.t. |x - y| = r \right\}.$$
Thanks to this result, both constants $R_0, R_1$ are finite. For $r \geq R_1$, condition (B.4) is satisfied since $k$ is sufficiently positive. It is then enough to assume condition (B.6) holds on the open interval $(0, R_1)$. Under this assumption,

$$g(R_1) \leq 1 - \frac{\alpha c}{2} \int_0^{R_1} f(s)\varphi^{-1}(s)ds \leq 1 - \frac{\alpha c}{4} \int_0^{R_1} \Phi(s)\varphi^{-1}(s)ds. \quad (B.7)$$

Condition (B.7), in turn, is satisfied if

$$\alpha c \leq \frac{2}{\int_0^{R_1} \Phi(s)\varphi^{-1}(s)ds}.$$

So, by choosing, for $r < R_1$,

$$g'(r) = -\frac{\Phi(r)}{2\varphi(r)} \int_0^{R_1} \Phi(s)\frac{\varphi^{-1}(s)}{\varphi(s)}ds,$$

condition (B.6) is fulfilled if we choose the constant as

$$\alpha c = \frac{1}{\int_0^{R_1} \Phi(s)\varphi^{-1}(s)ds}.$$

At this point, we can show that the quantity $\Gamma$ is smaller than $-cf(r)$, with our choices of $f$ and $c$. Consider the scenario $r < R_1$. Then, we have (see [22, eq. (68)])

$$f''(r) \leq \frac{1}{4}rk(r)f'(r) - \frac{1}{2}f(r) \int_0^{R_1} \Phi(s)\varphi^{-1}(s)ds. \quad (B.8)$$

Consider, now, the scenario $r \geq R_0$. Then,

$$f'(r) = \frac{\varphi(r)}{2} = \frac{\varphi(R_0)}{2},$$

and $k(r)R_1(R_1 - R_0) \geq 8$ by construction of $R_1$. Moreover, we know that $r \geq R_0$, the function $\varphi$ is constant and, therefore $\Phi(r) = \Phi(R_0) + (r - R_0)\varphi(R_0)$. Also,

$$\int_0^{R_1} \Phi(s)\varphi^{-1}(s)ds \geq (R_1 - R_0)\Phi(R_1)\varphi^{-1}(R_0)/2.$$

This implies (see [22, eq. (69)])

$$f''(r) - \frac{1}{4}rk(r)f'(r) \leq -\frac{1}{2}f(r) \int_0^{R_1} \Phi(s)\varphi^{-1}(s)ds. \quad (B.9)$$

Putting together equations (B.8) and (B.9), we conclude the key relationship

$$\Gamma(t) \leq -cf(r(t))$$

at all times $t < T$. For any coupling $\gamma_t$ of the process $(x(t), y(t))$, we take the expectation on both sides of (B.2) and obtain

$$\mathbb{E}^{\gamma_t}[f(r(t))] = \mathbb{E}^{\gamma_t}[f(r(s))] + \int_s^t \mathbb{E}^{\gamma_u}[m(u)]du \quad (B.10)$$

32
for any $s \leq t < T$. Let $\Upsilon(t) := \mathbb{E}^\gamma_t[f(r(t))]$. Then, differentiating (B.10) with respect to time yields

$$\Upsilon'(t) = \mathbb{E}^\gamma_t[m(t)].$$

Since we have proved that $m(t) \leq \Gamma(t) \leq -cf(r(t))$ for $t < T$, we deduce

$$\Upsilon'(t) \leq -c\Upsilon(t).$$ (B.11)

Thanks to standard Gronwall’s lemma, we deduce

$$\Upsilon(t) \leq \Upsilon(s)e^{-c(t-s)}$$ (B.12)

for all $s \leq t \leq T$. Hence, $t \to e^{ct}\mathbb{E}^\gamma_t[d_f(x(t), y(t))]$ is a decreasing function of time. This key result implies

$$W_f(\mu_{t,\beta}, \nu_{t,\beta}) \leq \mathbb{E}^{\gamma_{t,\beta}}[d_f(x(t), y(t))] \leq e^{-ct}\mathbb{E}^{\gamma_{t,\beta}}[d_f(x_0, y_0)],$$

where $\mu_{t,\beta}$ and $\nu_{t,\beta}$ denote the time-$t$ evolved probability measures of the process $x(t)$ with respect to the initial distributions $\mu$ and $\nu$ respectively, and $\gamma_{t,\beta}$ denotes their coupling, given a realization of $\beta$. Taking the infimum over all couplings $\gamma_{t,\beta}$, we conclude. \hfill \Box

**Proof of Proposition 3.3** By construction, the function $f$ in Proposition 3.2 is concave, increasing and satisfies $f(0) = 1, f'(0) = 1$. This implies that $f'(x)x \leq f(x) \leq x$. Moreover, $\frac{\varphi(R_0)}{2} \leq f' \leq 1$ thanks to the properties of $\varphi$ and $g$. Hence,

$$\frac{\varphi(R_0)}{2}\|x - y\| \leq d_f(x, y) \leq \|x - y\|$$

for any $x, y \in \mathbb{R}^d$. For any coupling $\gamma_{t,\beta}$ of $\mu_{t,\beta}$ and $\nu_{t,\beta}$,

$$\frac{\varphi(R_0)}{2}\mathbb{E}^{\gamma_{t,\beta}}[\|x(t) - y(t)\|] \leq \mathbb{E}^{\gamma_{t,\beta}}[d_f(x(t), y(t))] \leq e^{-ct}\mathbb{E}^{\gamma_{t,\beta}}[d_f(x_0, y_0)] \leq e^{-ct}\mathbb{E}^{\gamma_{t,\beta}}[\|x_0 - y_0\|].$$

Let $K := 2\varphi(R_0)^{-1}$. Taking the infimum over all couplings $\gamma_{t,\beta}$ yields

$$KW^1(\mu_{t,\beta}, \mu_{t+r,\beta}) \leq W_f(\mu_{t,\beta}, \mu_{t+r,\beta})$$

for all $t > 0$. Hence, if $\mu_{t,\beta})_{t \geq 0}$ is a Cauchy sequence with respect to $W_f$, it will be a Cauchy sequence with respect to $W^1$ as well. Moreover, with $\mu_{t,\beta}$ denoting the Lebesgue density of the measure $\mu_{t,\beta}$, the Hardy-Landau-Littlewood inequality [6, 7, 9] entails

$$\|u_{t,\beta} - u_{t,\beta}\|_{L^1} \leq C\|\nabla(u_{t,\beta} - u_{t,\beta})\|_{L^1} W^1(\mu_{t,\beta}, \mu_{t+r,\beta})$$

for some constant $C > 0$. Since the $L^1$ norm of the gradient is bounded (see Appendix A, Lemma [A.1]), we have

$$\|u_{t,\beta} - u_{t,\beta}\|_{L^1} \leq C W^1(\mu_{t,\beta}, \mu_{t+r,\beta})$$

for some constant $\tilde{C} > 0$. Hence, $(p_{t,\beta})_{t \geq 0}$ is a Cauchy sequence in $L^1$. Since $L^1$ is complete, the sequence converges in $L^1$, that is, for any initial condition $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$p_\beta = \lim_{t \to \infty} \Phi(t, \beta)p_\mu \in L^1.$$

\hfill \Box

33
C Proofs of results in Section 4

Proof of Proposition 4.1. In the stochastic setting, the non-autonomous Fokker-Planck equation (2.2) naturally extends to a random Fokker-Planck equation in terms of common noise sample paths $\beta$, which we here write in compact form as

$$\partial_t q = F(\theta_t \beta, q) \tag{C.1}$$

for some appropriate functional $F$. Subsequently, in analogy to the discussion in Section 2.1, the stochastic Fokker Planck equation (1.3) is obtained via the transformation $y = x - \eta \beta$, yielding the analogous form when choosing the stochastic integral to be of Stratonovich type. The cocycle property is obtained from the existence and uniqueness of solutions of (2.2) for almost all sample paths, as established in Section 2. The evolution operator $\Phi$ of the random Fokker Planck equation (C.1) is given by

$$\Phi(t, \beta, q) = q + \int_0^t F(\theta_s \beta, \Phi(s, \beta, q))ds.$$ 

Following closely the argument in Arnold [2, Proof of Theorem 2.2.1], we prove the cocycle property (for almost all $\beta \in \Omega_B$). Let $s, t \in \mathbb{R}$ and assume $s > 0, t > 0$ (the remaining cases are analogous). Then,

$$\Phi(t, \theta_s \beta, \Phi(s, \beta, q)) = \Phi(s, \beta, q) + \int_0^t F(\theta_{u+s} \beta, \Phi(u, \theta_s \beta, \Phi(s, \beta, q)))du$$

$$= q + \int_0^t F(\theta_s \beta, \Phi(s, \beta, q))ds$$

$$+ \int_s^{t+s} F(\theta_z \beta, \Phi(z-s, \theta_s \beta, \Phi(s, \beta, q)))dz,$$

where $z = u + s$. Therefore, the function

$$\tilde{\Phi}(u, \beta, q) := \begin{cases} 
\Phi(u, \beta, q) & \text{if } 0 \leq u \leq s \\
\Phi(u-s, \theta_s \beta, \Phi(s, \beta, q)) & \text{if } s \leq u \leq s+t
\end{cases}$$

satisfies

$$\tilde{\Phi}(t+s, \beta, q) = q + \int_0^{t+s} F(\theta_s \beta, \tilde{\Phi}(u, \beta, q))du.$$ 

By uniqueness, for $\mathbb{P}$-a.e. $\beta \in \Omega_B$,

$$\Phi(t+s, \beta, q) = \tilde{\Phi}(t+s, \beta, q) = \Phi(t, \theta_s \beta, \tilde{\Phi}(s, \beta, q)).$$

Proof of Theorem 4.2. From proposition 3.2 we deduce that there exist a constant $c > 0$ and an increasing and convex function $f$ such that for any $t > 0, \beta \in C^{1/2}(\mathbb{R}, \mathbb{R}^d)$ and initial probability measures $\mu, \nu \in \mathcal{M}(\mathbb{R}^d)$,

$$W_f(\mu, \nu, \beta) \leq e^{-ct} W_f(\mu, \nu).$$
with $\mu_{t,\beta} := \Psi(t, \beta)\mu$ and similarly for $\nu_{t,\beta}$, where $\Psi$ denotes the time-$t$ evolution operator for the measure $\mu$, associated to the time-$t$ evolution operator $\Phi$ of the random Fokker Planck equation. We show that $(\mu_{t,\beta})_{t>0}$ is a Cauchy sequence in a pullback sense with respect to the $W_f$ metric, that is, $\forall \epsilon > 0 \exists \tau > 0$
\[ W_f (\Psi(t, \theta_{-\tau}\beta)\mu, \Psi(t + \tau, \theta_{-(t+\tau)}\beta)\mu) < \epsilon. \]

Exploiting the pullback operator and the fact that we have a contraction, we deduce
\[ W_f (\Psi(t, \theta_{-\tau}\beta)\mu, \Psi(t + \tau, \theta_{-(t+\tau)}\beta)\mu) = W_f (\Psi(t, \theta_{-\tau}\beta)\mu, \Psi(t, \theta_{-\tau}\beta) \circ \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu) \leq e^{-ct} W_f (\mu, \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu). \]

Then, since $f$ is concave and increasing by construction,
\[ W_f (\mu, \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu) \leq f(W^1(\mu, \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu)) \leq W^1(\mu, \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu). \]

We observe that $C^{1/2}(\mathbb{R}, \mathbb{R}^d)$ is a subset of full Wiener measure $\mathbb{P}_\beta$ of the sample path space $\Omega_B$. Taking the expectation $E^\beta$ with respect to $\mathbb{P}_\beta$ implies
\[ E^\beta [W_f (\mu, \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu]] \leq E^\beta [W^1(\mu, \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu)]. \]

Let $p_\mu$ and $p_{\tau, t+\tau\beta}$ denote the Lebesgue densities of $\mu$ and $\Psi(\tau, \theta_{-(t+\tau)}\beta)\mu$ respectively. Their product will be the density of the product measure, which is a simple example of a coupling measure. Therefore,
\[ E^\beta [W^1(\mu, \Psi(\tau, \theta_{-(t+\tau)}\beta)\mu)] \leq E^\beta \left[ \int_{\mathbb{R}^{2d}} \|x - y\| p_{\tau, t+\tau\beta}(x) p_\mu(y) dx dy \right] \]
\[ = \int_{\mathbb{R}^{2d}} \|x - y\| E^\beta [p_{\tau, t+\tau\beta}(x)] p_\mu(y) dx dy. \]

Let us define $p_\tau := E^\beta [p_{\tau, t+\tau\beta}]$. We notice that this expectation does not depend on $t$ since we are integrating over all $\Omega_B$ and $\theta_{-(t+\tau)}\beta = \theta_{-\tau}\tilde{\beta}$ for some $\tilde{\beta} \in \Omega_B$. Then
\[ \int_{\mathbb{R}^{2d}} \|x - y\| p_\tau(x) p_\mu(y) dx dy = \int_{\mathbb{R}^d} p_\mu(y) \int_{\mathbb{R}^d} \|x - y\| p_\tau(x) dx dy \]
\[ \leq \int_{\mathbb{R}^d} p_\mu(y) \left[ \int_{\mathbb{R}^d} \|x\| p_\tau(x) + \|y\| p_\tau(x) dx \right] dy \]
\[ = \int_{\mathbb{R}^d} p_\mu(y) \left( \int_{\mathbb{R}^d} \|x\| p_\tau(x) dx + \|y\| \right) dy \]
\[ = \int_{\mathbb{R}^d} p_\tau(x) \|x\| dx + \int_{\mathbb{R}^d} p_\mu(y) \|y\| dy. \]

The second term on the RHS of the equation above is bounded. For what concerns the first term, we further notice that $p_\tau$ is the forward solution at time $t = \tau$, with initial condition $p_\mu$ at time $t = 0$, of the autonomous Fokker Planck equation for the SDE
\[ \frac{\partial}{\partial t} p = \Delta V(x)p + \nabla V(x) \frac{\partial p}{\partial x} + \frac{1}{2} (\sigma^2 + \eta^2) \frac{\partial^2 p}{\partial x^2}. \]
Applying the results from Sections 2 and 3 to the autonomous setting, we deduce this equation admits a unique attractor and, in particular,

$$\lim_{t \to \infty} p_t = p_\rho \quad \text{in } L^1,$$

where $p_\rho$ denotes the density of the stationary measure. In fact, we observe that the fixed point $p_\rho$ is invariant under the autonomous evolution operator $\tilde{\Phi} = \Phi(\cdot, 0)$, i.e.

$$\tilde{\Phi} p_\rho = p_\rho.$$

Thanks to the discussion in Appendix A, we deduce that $\tilde{\Phi}$ maps $L^1$ functions into the Schwartz space $S$ of rapidly decreasing functions. Therefore, $p_t$ converges exponentially fast to $p_\rho$ as $t \to \infty$ in $S$. Next, we consider

$$\int_{\mathbb{R}^d} p_\tau(x) \|x\| \, dx = \int_{\mathbb{R}^d} (p_\tau(x) - p_\rho(x)) \|x\| \, dx + \int_{\mathbb{R}^d} p_\rho(x) \|x\| \, dx$$

$$\leq \int_{\mathbb{R}^d} |p_\tau(x) - p_\rho(x)| \|x\| \, dx + \int_{\mathbb{R}^d} p_\rho(x) \|x\| \, dx.$$

For any $\epsilon > 0$, there exists $T > 0$ such that for any $\tau > T$,

$$\int_{\mathbb{R}^d} |p_\tau(x) - p_\rho(x)| \|x\| \, dx < \epsilon$$

Let us define

$$C := \sup_{\tau \in [0, T]} \int_{\mathbb{R}^d} p_\tau(x) \|x\| \, dx < \infty,$$

which is finite since $\int_{\mathbb{R}^d} p_\tau(x) \|x\| \, dx$ is finite for any $\tau$ and the supremum is taken over a finite time interval. Then,

$$\int_{\mathbb{R}^d} p_\tau(x) \|x\| \, dx < 1_{\tau \leq T} C + 1_{\tau > T} \left( \epsilon + \int_{\mathbb{R}^d} p_\rho(x) \|x\| \, dx \right)$$

$$\leq \max \left\{ C, \epsilon + \int_{\mathbb{R}^d} p_\rho(x) \|x\| \, dx \right\}.$$ 

Consequently, there exists a constant $D < \infty$ such that for any $\tau > 0$

$$\mathbb{P}^{\beta}[W_f(\mu, \Psi(\tau, \theta_{-(t+\tau)} \beta) \mu)] < D,$$

hence $W_f(\mu, \Psi(\tau, \theta_{-(t+\tau)} \beta) \mu) < D$ for $\mathbb{P}_B$-almost all $\beta \in \Omega_B$. This shows that

$$\lim_{t \to \infty} W_f \left( \Psi(t, \theta_{-t} \beta) \mu, \Psi(t + \tau, \theta_{-(t+\tau)} \beta) \mu \right) = 0$$

for all $\tau > 0$ and $\mathbb{P}_B$-almost all $\beta \in \Omega_B$. Therefore, $(\Psi(t, \theta_{-t} \beta) \mu)_{t > 0}$ is $\mathbb{P}_B$-almost surely a Cauchy sequence with respect to the $W_f$ metric. Let $p_{t, \beta}$ denote the Lebesgue density of $\Psi(t, \theta_{-t} \beta) \mu$. Thanks to Proposition 3.3

$$\lim_{t \to \infty} p_{t, \beta} = p_\beta \in L^1$$
for all $t > 0$ and $\mathbb{P}_B$-almost all $\beta \in \Omega_B$. Finally, we remark that the limit point $p_\beta \in L^1$ is invariant under the pullback flow. We have

$$p_\beta = \Phi(t, \beta)p_{\theta_t \beta}$$

for all $t > 0$. In light of the discussion in Appendix A, $\Phi$ maps $L^1$ functions to $S$ functions. In other words, $p_\beta \in S$. We conclude that for $\mathbb{P}_B$-almost all $\beta \in \Omega_B$ and initial probability density $p \in L^1$, there exists a unique pullback attractor for (2.2)

$$p_\beta = \lim_{t \to \infty} \Phi(t, \theta_t \beta)p \in S.$$

Proof of Proposition 4.4. By construction, for every $C \in \mathcal{B}(\mathbb{R}^d)$, any Borel measure $\nu$ on $\mathbb{R}^d$ with Lebesgue density $p_\nu \in S$ and for $\mathbb{P}_B$-almost all $\beta \in \Omega_B$, we have for all $s \leq t$

$$\nu_t(C) := \int_{\Omega_B} \phi_s(t - s; \theta_s \omega, \theta_s \beta)p(C)\mathbb{P}_B(d\omega) = \int_{C} \Phi(t - s; \theta_s \beta)p_{\nu_s}dx = \int_{C} p_{\nu_t}dx, \quad (C.2)$$

and, using the results from proposition 4.3 and (C.2),

$$\mu_\beta(C) = \int_{\Omega_B} \mu_{\omega, \beta}(C)p(C)\mathbb{P}_B(d\omega) = \int_{\Omega_B} \lim_{\tau \to \infty} \phi(\tau, \theta_{-\tau} \omega, \theta_{-\tau} \beta)p(C)\mathbb{P}_B(d\omega)$$

$$= \lim_{\tau \to \infty} \int_{\Omega_B} \phi(\tau, \theta_{-\tau} \omega, \theta_{-\tau} \beta)p(C)\mathbb{P}_B(d\omega) = \lim_{\tau \to \infty} \int_{C} \Phi(\tau, \theta_{-\tau} \beta)p_{\rho}dx$$

$$= \int_{C} \lim_{\tau \to \infty} \Phi(\tau, \theta_{-\tau} \beta)p_{\rho}dx$$

by which the result follows from the fact that $\mu_\beta(C) = \int_{C} p_\beta dx$, for all $C \in \mathcal{B}(\mathbb{R}^d)$. \qed

Proof of Proposition 4.5. The Dirac measure $\delta_{p_\beta}$ is the disintegration of a Markov measure of the random dynamical system $\Phi$ on $\Omega_B \times S$, associated to the stochastic Fokker-Planck (1.3). The $S$-marginal of this Markov measure

$$P := \int_{\Omega_B} \delta_{p_\beta}\mathbb{P}_B(d\beta) \quad (C.3)$$

is the corresponding stationary measure of (1.3). Application of Birkhoff’s Ergodic Theorem then yields that time-averages of $(\mathbb{P}^\top)$-integrable observables $g : S \to \mathbb{R}$ satisfy

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau g(\Phi(\tau, \beta)p)dt = \int_A g(p)\mathbb{P}(dp) = \int_{\Omega_B} g(p_\beta)\mathbb{P}_B(d\beta),$$

$\mathbb{P}_B \times \mathbb{P}$-almost surely. By Elton’s Ergodic Theorem [24], this relation holds in fact $\mathbb{P}_B$-almost surely if $g$ is continuous. \qed

37
D Exact solutions for the Ornstein-Uhlenbeck SDE with intrinsic and common additive noise

In this appendix, we present the calculations of the closed expressions for common noise pullback attractors of the one-dimensional Ornstein-Uhlenbeck SDE with additive intrinsic and common noise, discussed in Section 1.3.1. The flow of the SDE (1.7) from time $s$ to $t$ for fixed noise realisations $\beta$ and $\omega$ is explicitly given by

$$\phi(t - s, \theta_s \omega, \theta_s \beta)x(s) = x(s)e^{-a(t-s)} + \eta \int_s^t e^{-a(t-u)}d\beta(u) + \sigma \int_s^t e^{-a(t-u)}d\omega(u).$$

Averaging this equation over the intrinsic noise yields

$$\int_{\Omega_W} \phi(t - s, \theta_s \omega, \theta_s \beta)x(s)\mathbb{P}_W(d\omega) = x(s)e^{-a(t-s)} + \eta \int_s^t e^{-a(t-u)}d\beta(u) + \sigma \int_s^t e^{-a(t-u)}dW(u).$$

It’s important to emphasize that the integral with respect to the single path $\beta$ is a real number while the integral with respect to the intrinsic noise $W$ is a Gaussian distribution. The density of the distribution in (D.1) is given by

$$p(x,t) = \sqrt{\frac{a}{\pi \sigma^2 (1 - e^{-2a(t-s)})}} \exp \left(-\frac{a}{\sigma^2 (1 - e^{-2a(t-s)})} (x - m_\beta(t,s))^2 \right),$$

where

$$m_\beta(t,s) := x(s)e^{-a(t-s)} + \eta \int_s^t e^{-a(t-u)}d\beta(u).$$

It is readily checked that indeed the density (D.2) is a solution of the stochastic Fokker-Planck equation (1.3) with $V(x) = \frac{a}{2}x^2$.

Averaging (D.1) over the common noise yields

$$\int_{\Omega_B} \int_{\Omega_W} \phi(t - s, \theta_s \omega, \theta_s \beta)x(s)\mathbb{P}_W(d\omega)\mathbb{P}_B(d\beta) = x(s)e^{-a(t-s)} + \eta \int_s^t e^{-a(t-u)}dB(u) + \sigma \int_s^t e^{-a(t-u)}dW(u),$$

where now both integrals represent Gaussian distributions. The density of the distribution (D.3) is

$$\tilde{p}(x,t) = \int_{\Omega_B} p(x,t)\mathbb{P}_B(d\beta) = \sqrt{\frac{a}{\pi (\eta^2 + \sigma^2)(1 - e^{-2a(t-s)})}} \exp \left(-\frac{a}{(\eta^2 + \sigma^2)(1 - e^{-2a(t-s)})} x^2 \right),$$

which in turn is a solution of the Fokker-Planck equation of the SDE (1.1) with $V(x) = \frac{a}{2}x^2$: $\frac{\partial \tilde{p}}{\partial t} = a \tilde{p} + ax \frac{\partial \tilde{p}}{\partial x} + \frac{1}{2} (\sigma^2 + \eta^2) \frac{\partial^2 \tilde{p}}{\partial x^2}$. 

38
We conclude this section with a discussion on pullback attractors. The pullback attractor of the SDE (1.7) with respect to both intrinsic and common noise is

$$\alpha(\omega, \beta) := \lim_{s \to -\infty} \phi(t-s, \theta_s \omega, \theta_s \beta) x(s) = \eta \int_{-\infty}^{0} e^{au} d\beta(u) + \sigma \int_{-\infty}^{0} e^{au} d\omega(u).$$

(D.5)

This is a point attractor, confirming that, at the SDE level, the system is synchronizing. The fiberwise measures resulting from disintegration (see Section 4) are therefore

$$\mu_{\omega, \beta} = \delta_{\alpha(\omega, \beta)}.$$

Integrating with respect to the intrinsic noise yields

$$\mu_\beta := \int_{\Omega_W} \mu_{\omega, \beta} \mathbb{P}_W (d\omega) = \eta \int_{-\infty}^{0} e^{au} d\beta(u) + \sigma \int_{-\infty}^{0} e^{au} dW(u).$$

This is normally distributed with variance depending on the intensity of the intrinsic noise $\sigma$ and mean depending on the intensity of the common noise $\eta$. Its density is

$$p_\beta(x) = \sqrt{\frac{a}{\pi \sigma^2}} \exp \left\{ -\frac{a}{\sigma^2} \left( x - \eta \int_{-\infty}^{0} e^{au} d\beta(u) \right)^2 \right\}.$$

(D.6)

Finally, integrating over all common noise realizations we obtain the stationary measure

$$\rho = \int_{\Omega_B} \mu_\beta \mathbb{P}_B (d\beta) = \int_{\Omega_B} \int_{\Omega_W} \mu_{\omega, \beta} \mathbb{P}_W (d\omega) \mathbb{P}_B (d\beta)$$

$$= \eta \int_{-\infty}^{0} e^{au} dB(u) + \sigma \int_{-\infty}^{0} e^{au} dW(u),$$

with density

$$p_\rho(x) = \sqrt{\frac{a}{\pi (\sigma^2 + \eta^2)}} \exp \left\{ -\frac{a}{(\eta^2 + \sigma^2)} x^2 \right\}.$$

Of course, we also have $p_\rho(x) = \int_{\Omega_B} p_\beta(x) \mathbb{P}_B (d\beta)$ and $p_\rho(x) = \lim_{(t-s) \to \infty} \bar{p}(x, t)$ confirming global convergence of solutions of (D.5) to the stationary measure in forward and pullback sense.

**Acknowledgements**

We are grateful to Sergey Zelik for extended discussions and input on the material in Section 2.2 and Appendix A and to Martin Hairer for pointing out [22]. FG has been supported by the EPSRC Centre for Doctoral Training in Financial Computing and Analytics. JSWL acknowledges support by the UK Royal Society (NAF\R1\180236), the EU H2020 ITN CRITICS (643073) and the London Mathematical Laboratory (LML) through its External Fellowship scheme.
References

[1] L. F. Abbott, K. Rajan, and H. Sompolinsky, Interactions between intrinsic and stimulus-evoked activity in recurrent neural networks, The Dynamic brain: an exploration of neuronal variability and its functional significance (Ding M, Glanzman DL, eds), (2011), pp. 65–82.

[2] L. Arnold, Random Dynamical Systems, Springer, Berlin, 1998.

[3] L. Arnold and M. Scheutzow, Perfect cocycles through stochastic differential equations, Probability theory and related fields, 101 (1995), pp. 65–88.

[4] A. Bain and D. Crisan, A continuous time particle filter, in Fundamentals of Stochastic Filtering, Springer, 2009, pp. 221–256.

[5] P. W. Bates, K. Lu, and B. Wang, Random attractors for stochastic reaction–diffusion equations on unbounded domains, Journal of Differential Equations, 246 (2009), pp. 845–869.

[6] V. Bogachev and A. Shaposhnikov, Lower bounds for the Kantorovich distance, Doklady Mathematics, 91 (2015), pp. 91–93.

[7] V. Bogachev, F. Wang, and A. Shaposhnikov, Estimates of the Kantorovich norm on manifolds, Doklady Mathematics, 92 (2015), pp. 494–499.

[8] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, Fokker-Planck-Kolmogorov Equations, vol. 207, American Mathematical Soc., 2015.

[9] V. I. Bogachev, A. V. Shaposhnikov, and F.-Y. Wang, Sobolev–Kantorovich inequalities under CD(0,∞) condition, Communications in Contemporary Mathematics, (2021), p. 2150027.

[10] E. Boissard and T. Le Gouic, On the mean speed of convergence of empirical and occupation measures in Wasserstein distance, Annales de l’IHP Probabilités et Statistiques, 50 (2014), pp. 539–563.

[11] F. Bolley, I. Gentil, and A. Guillin, Convergence to equilibrium in wasserstein distance for fokker-planck equations, Journal of Functional Analysis, 263 (2012), pp. 2430–2457.

[12] P. C. Bressloff, Stochastic Fokker-Planck equation in random environments, Physical Review E, 94 (2016), p. 042129.

[13] T. Caraballo and P. Kloeden, The pathwise numerical approximation of stationary solutions of semilinear stochastic evolution equations, Applied Mathematics and Optimization, 54 (2006), pp. 401–415.

[14] R. Carmona and F. Delarue, Probabilistic Theory of Mean Field Games with Applications I-II, vol. 83-84 of Probability Theory and Stochastic Modelling, Springer, 2018.
[15] J. Cerdà, *Linear functional analysis*, vol. 116, American Mathematical Soc., 2010.

[16] P. P. Clément, H. J. Heijmans, S. Angenent, C. J. van Duijn, and B. de Pagter, *One-parameter semigroups*, CWI monographs, 5 (1987).

[17] H. Crauel and F. Flandoli, *Attractors for random dynamical systems*, Probability Theory and Related Fields, 100 (1994), pp. 365–393.

[18] ——, *Additive noise destroys a pitchfork bifurcation*, Journal of Dynamics and Differential Equations, 10 (1998), pp. 259–274.

[19] M. Del Giudice, S. Bo, S. Grigolon, and C. Bosia, *On the role of extrinsic noise in microRNA-mediated bimodal gene expression*, PLoS computational biology, 14 (2018), p. e1006063.

[20] G. Drótos, T. Bódai, and T. Tél, *On the importance of the convergence to climate attractors*, The European Physical Journal Special Topics, 226 (2017), pp. 2031–2038.

[21] J. Duan and W. Wei, *Effective dynamics of stochastic partial differential equations*, Elsevier, 2014.

[22] A. Eberle, *Reflection couplings and contraction rates for diffusions*, Probability theory and related fields, 166 (2016), pp. 851–886.

[23] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Stochastic gene expression in a single cell*, Science, 297 (2002), pp. 1183–1186.

[24] J. H. Elton, *An ergodic theorem for iterated maps*, Ergodic Theory and Dynamical Systems, 7 (1987), pp. 481–488.

[25] L. C. Evans, *Partial differential equations and Monge-Kantorovich mass transfer*, Current developments in mathematics, 1997 (1997), pp. 65–126.

[26] ——, *Partial differential equations*, American Mathematical Society, 1998.

[27] B. Fehrman and B. Gess, *Well-posedness of nonlinear diffusion equations with nonlinear, conservative noise*, Archive for Rational Mechanics and Analysis, 233 (2019), pp. 249–322.

[28] F. Flandoli and H. Lisei, *Stationary conjugation of flows for parabolic spdes with multiplicative noise and some applications*, Stochastic analysis and applications, 22 (2004), pp. 1385–1420.

[29] F. Flandoli and B. Schmalfuss, *Random attractors for the 3d stochastic navier-stokes equation with multiplicative white noise*, Stochastics: An International Journal of Probability and Stochastic Processes, 59 (1996), pp. 21–45.

[30] M. Ghil, M. D. Chekroun, and E. Simonnet, *Climate dynamics and fluid mechanics: Natural variability and related uncertainties*, Physica D: Nonlinear Phenomena, 237 (2008), pp. 2111–2126.
[31] M. B. Giles and C. Reisinger, *Stochastic finite differences and multilevel Monte Carlo for a class of spdes in finance*, SIAM Journal on Financial Mathematics, 3 (2012), pp. 572–592.

[32] F. Herrerías-Azué and T. Galla, *The effects of heterogeneity on stochastic cycles in epidemics*, Scientific reports, 7 (2017), p. 13008.

[33] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, vol. 113, springer, 2014.

[34] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, vol. 54 of Encyclopedia of mathematics and its applications, Cambridge University Press, 2nd ed., 1993.

[35] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, vol. 23, Springer-Verlag Berlin Heidelberg, 1992.

[36] P. E. Kloeden and M. Rasmussen, *Nonautonomous dynamical systems*, no. 176 in Mathematical Surveys and Monographs, American Mathematical Soc., 2011.

[37] T. G. Kurtz and J. Xiong, *Particle representations for a class of nonlinear spdes*, Stochastic Processes and their Applications, 83 (1999), pp. 103–126.

[38] X. Lang, Q. Lu, and J. Kurths, *Phase synchronization in noise-driven bursting neurons*, Physical Review E, 82 (2010), p. 021909.

[39] K. Lu and B. Wang, *Wong–zakai approximations and long term behavior of stochastic partial differential equations*, Journal of Dynamics and Differential Equations, 31 (2019), pp. 1341–1371.

[40] B. Øksendal, *Stochastic differential equations*, Springer, 2003.

[41] G. A. Pavliotis, *Stochastic processes and applications: diffusion processes, the Fokker-Planck and Langevin equations*, vol. 60 of Texts in Applied Mathematics, Springer, 2014.

[42] E. PDEs, *Measures and capacities*, From the Poisson equation to Nonlinear Thomas-Fermi problems, EMS Tracts in Mathematics, 23 (2016).

[43] A. Pikovsky, J. Kurths, M. Rosenblum, and J. Kurths, *Synchronization: a universal concept in nonlinear sciences*, no. 12 in Cambridge Nonlinear Science series, Cambridge University Press, 2003.

[44] F. Sagués, J. M. Sancho, and J. García-Ojalvo, *Spatiotemporal order out of noise*, Reviews of Modern Physics, 79 (2007), p. 829.

[45] P. S. Swain, M. B. Elowitz, and E. D. Siggia, *Intrinsic and extrinsic contributions to stochasticity in gene expression*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 12795–12800.

[46] P. Thomas, *Population growth affects intrinsic and extrinsic noise in gene expression*, bioRxiv, (2018), p. 362368.
[47] H. Triebel, Interpolation theory, function spaces, differential operators, Bull. Amer. Math. Soc.(NS), 2 (1977), pp. 339–345.

[48] V. S. Varadarajan, On the convergence of sample probability distributions, Sankhyā: The Indian Journal of Statistics (1933-1960), 19 (1958), pp. 23–26.

[49] M.-K. von Renesse and K.-T. Sturm, Transport inequalities, gradient estimates, entropy and ricci curvature, Communications on pure and applied mathematics, 58 (2005), pp. 923–940.

[50] Z. Wang and S. Zhou, Random attractor for stochastic reaction–diffusion equation with multiplicative noise on unbounded domains, Journal of Mathematical Analysis and Applications, 384 (2011), pp. 160–172.

[51] X. Yan, X. Liu, and M. Yang, Random attractors of stochastic partial differential equations: A smooth approximation approach, Stochastic Analysis and Applications, 35 (2017), pp. 1007–1029.